DOCUMENT RESUME

ED 215 867 SE 037 055

AUTHOR Rogerson, Alan, Ed.
TITLE Co-operation Between Science Teachers and Mathematics Teachers. Volumes 1-6.
INSTITUTION International Council of Scientific Unions, Paris (France).
SPONS AGENCY United Nations Educational, Scientific, and Cultural Organization, Paris (France).
PUB DATE [78]
NOTE 219p.
AVAILABLE FROM John Goodman & Sons (Printers) Ltd., Cardigan Street, Birmingham, England B4 7RS

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Biology; Calculus; Chemistry; Geography; Inservice Teacher Education; Instructional Materials; Interdisciplinary Approach; *Mathematical Applications; Mathematical Models; *Mathematics Instruction; Models; Physics; *Science Instruction; Secondary Education; *Secondary School Mathematics; *Secondary School Science; *Teacher Education

ABSTRACT

This document contains six separate works, titled: (1) Functions and Physics; (2) Links Between Geography and Mathematics; (3) Our Inheritance: Common Ground for the Mathematics and Biology Teacher; (4) Mathematics and Chemistry: The Classroom Interface; (5) Mathematical Modeling; and (6) Mathematical Modeling with Calculus. This series of booklets arose out of a conference held in Bielefeld, West Germany, in 1978, which brought together teachers of mathematics and science. The materials are geared towards providing common ground for cooperation between mathematicians and scientists, through pointing out areas where interests overlap. While the presentation is geared largely towards teacher education, much is presented with a view towards use in mathematics and science classrooms, primarily in the secondary schools. (MP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Introduction

This booklet is mainly directed towards mathematics and physics teachers, and presents three areas of overlap where the central mathematical idea of *function* can be linked to Hooke's Law, acceleration and electrical resistors respectively.

What does cooperation between mathematics and physics *mean*? It could mean practical liaison within a school on matters of time-tabling, team teaching or curricular content. This presupposes that teachers are willing to cooperate (the vital factor) and this booklet presents some particular content areas where cooperation can take place. This could take the form of informal discussion about consistent terminology, a deeper awareness of each subject's approach and even the use of experiments in mathematics.

There are, of course, many areas that mathematics and physics have in common and it is hoped that these three particular examples will encourage a wider personal cooperation between mathematics and physics teachers in schools. Through the teachers, it is our pupils who will be the chief beneficiaries of this cooperation.

Part 1

Hooke's Law and the idea of a function

A Physicist and a Mathematician take a close look at an elementary topic.

Robert Morris
Charles Taylor

1. Introduction

1.1 Perspectives on physics and mathematics are constantly changing. Older ways of teaching give way to newer ones in an evolutionary sequence. But in recent years, the mutations in both subjects have tended to be more frequent and the pace of change has increased. It is hard enough, these days, for a teacher to keep up-to-date with his own subject; to expect him to keep up with contemporary thought on the teaching of *two* subjects is to ask too much. Yet mathematics and physics are so interdependent that teachers of the two subjects must, at least, be able to talk to each other and each must be aware of the approach being made by the other.

1.2 This maxim is especially important for those who teach the younger age groups. With younger children mathematical concepts have their roots in the physical world. Children should meet a new idea in its concrete form before they can formalize and, later, abstract it. But there is a converse. The young child needs also to be able to talk about and write about each new physical idea. For this he needs symbols, mathematical as well as linguistic. In this sense, the mathematical symbols he uses are there for reasons of convenience and economy of expression.

1.3 So, hard though it may be, it is necessary for teachers of mathematics to absorb something of the spirit of the contemporary approach to the teaching of physics, and for teachers of physics to try to absorb some of the newer words and symbols that have come into use in the mathematics lesson. How can this be done? Our plan is to take a close look at a sequence of well-established, but still relevant, topics which will be familiar to both mathematics and physics teachers and to consider sympathetically the changes that have taken place in their treatment. The first example we have chosen is that of Hooke's Law. It illustrates a very simple mathematical relationship that should, ideally, be reinforced by teaching it both in the physics laboratory and in the mathematics room. But clearly it is essential that the children are helped to recognize that it *is* the same topic.

2. Presentation

2.1 In the past, the aim of the physics teacher was to get to "the result" as quickly as possible. Exceptions to the rule were ignored and deviations from the expected result were brushed aside. On the rare, but golden occasions when the experiment was demonstrated to the class, the objects to be experimented upon (the coiled spring, the steel wire), and the limits of the variables, were carefully chosen to ensure that the simple form of the law was confirmed. The target was to arrive at the relationship

$$W = kx$$

in the shortest possible time and by the most reliable route [W is the suspended weight and x the resultant extension in a particular experiment and k is a constant]. The results (Table 1) were unmistakeable. Few students would be unable to "guess" the law, or to see how the results fitted (or "proved") the expected law.

W;	Suspended weight in kilograms:	1	2	3	4	5	6
x;	Extension of spring in mm:	20	40	60	80	100	120

Table 1

The teacher would then explain that the gravitational pull on the weights is equal and opposite to the tension in the material, which is the quantity that appears in the more usual formula, and so we can re-write the table:-

T:	Tension in kg. wt:	–1	–2	–3	–4	–5	–6
x;	Extension of spring in mm:	20	40	60	80	100	120

Table 2

2.2 The contemporary approach is quite different. The physics teacher knows that to lead the class in this way would be to mislead it. The real world is far less tidy, and the general aim is not exactly to disillusion children, but at least help them see that the neat behaviour of the spring or wire is only one possible type of behaviour. It may be chosen deliberately out of many other possibilities because it leads to results that are within the mathematical capability of the class to interpret. It is always important to avoid the notion that the spring or wire is behaving "ideally" and that all the other materials tried out by the class - cotton thread, nylon fishing line, lead wire, Sellotape, a piece of carrot etc. - are in some way struggling to achieve this ideal behaviour and failing. This is the real world and our ultimate selection is made with our own limitations in theoretical ability clearly in mind.

2.3 Furthermore in many physics laboratories the experiments will be done by the children themselves, each small group using a different material. The experiment is not done with a given law in mind but rather to "explore the relationship". In these circumstances the results cannot be 'cooked' and the many sources of deviation - the nature of the material, errors of observation, careless errors etc. - make it unlikely that an obvious relationship will be found. Even a set of results from a good Hookean object made by an observer who is unaware of Hooke's Law may not immediately enable the relationship to be guessed (Table 3).

T;	Tension in kgs. wt:	-1	-2	-3	-4	-5	-6
x;	Extension of spring in mm:	19	42	61	79	102	121

Table 3

2.4 The relationship may not be obvious; indeed for many of the test objects it may turn out that there is no relationship at all that is capable of being written in a manageable algebraic form.

Nevertheless, from the point of view of contemporary mathematics, *all* the sets of results may be said to define a function.

Thus the modern physicists' approach (with its stress on the real world) and the modern mathematician's approach (with its stress on the generality of the concept of a function) may, if not carefully handled, lead to considerable confusion!

3. Measures and numbers

3.1 Generally speaking, a physicist is concerned with number only as the magnitude of a measurement. A mathematician is concerned with number as an abstraction from the real world. Thus, for example, a physicist might be considering quantities such as 2 kg, 2 m, or 2 s; the mathematician would see that there is something in common between the three items, namely the abstract idea of the number 2. The same idea is also shared with his eyes, his ears and his feet, but not with his nose or fingers.

3.2 It is important to recognize a difference between a measure and a count. We use numbers to count eyes, sheep, buttons etc. and physicists and mathematicians would be unlikely to disagree on a matter such as the number of legs on a camel. But they might well see differently the measure of its height, 2 m. The mathematician is likely to think of this as in Fig. 1(a), whereas the physicist may recognize this as 2 m ± 0.5 m (Fig. 1(b)) or 200 cm ± 0.5 cm (Fig. 1(c)).

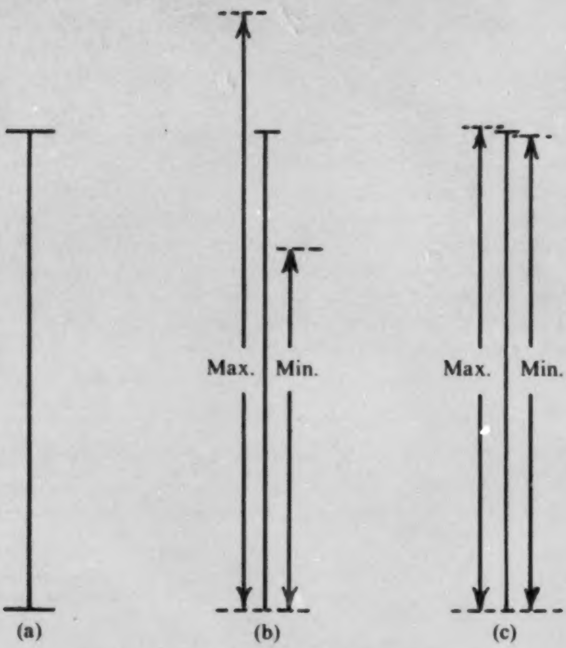


Fig. 1

3.3 These differences of approach are not trivial; they influence the formulation of relationships. Take for example, the formulation of Hooke's Law. The same formula

$$T = -kx$$

can now be seen to mean different things to contemporary physicists and mathematicians. To the physicist 'T' and 'x' are measured quantities and each symbol must have, associated with it, two features, a number and a unit. T (as in Table 3) may range from -1 kg to -6 kg; x may range from 19 mm to 121 mm. We hope that the pupils will eventually discover that k is about 0.5 – but what unit should be associated with this number? We clearly cannot equate kg weight with millimetres, and so k must have its own units of kg wt per mm.

In the abstract world of numbers, the product of two numbers is another number and so, to a mathematician, 'T' is the number you get by multiplying the number represented by 'k' by the number represented by 'x'.

4. Different kinds of numbers

4.1 Teachers of physics are so used to the idea of measuring *continuous* quantities [i.e. to measuring rather than counting] that they tend always to think of numbers as real. They may not be aware that children are sometimes nowadays taught to distinguish between several different number systems:-

Counting numbers	1, 2, 3, ----------
Positive rational numbers	-- $\frac{2}{3}$ --- $\frac{7}{8}$ --- $1\frac{1}{5}$
Integers	-- -2, -1, 0, 1, 2, 3, ------
Rational numbers	--- $-2\frac{3}{7}$---- $-\frac{1}{2}$----0---$\frac{1}{4}$---$7\frac{3}{5}$----
Real numbers	(-5.47 --) -- 0 -- (1.836 ---) ----

4.2 It is rather unlikely that an eighth grade physics class listening to the teacher's exposition of Hooke's Law will ask in what system of numbers the physics teacher is working! But if the question *were* put the physics teacher ought to realize that it is a fair one. He should therefore control his emotions and reassuringly explain that, in physics, we are concerned with the real world and that the field of real numbers is therefore the appropriate one in which to work.

5. The analysis of the results

5.1 At some point in the discussion the physics teacher will ask the question in relation to Table 3 "Do you notice anything?" The answer from a class that has met modern mathematics might well be 'a function'. The children will have met 'sets' and 'mappings' in their maths lessons and may therefore recognize the magnitudes of T in the table as a set of numbers. They will also think of the magnitudes of x as another set. They will further perceive that each member of the first set "maps on" to one, and only one, member of the second set. There is thus a one-to-one correspondence between the members of the two sets and this is enough for them to define a function.

5.2 It is sometimes, but not always, possible to wrap up a functional relationship into a neat little package. For example the relationship of Table 2 can be expressed as

$$T = -0.05\ x$$

and the pupils should also recognize that T and x do not merely represent the values in the table, but also all the other possible corresponding pairs of measures of T and x.

5.3 However, the results for the nylon fishing line, or the carrot, or even the approximate results of Table 3, are not capable of being packaged in a straightforward mathematical way. Nevertheless, from a mathematical standpoint, the tables of correspondence for all the different materials tested each pass the test which qualifies them to be deemed functional relationships.

5.4 In the physics lab – whether 'contemporary' or 'trad' – the relationship between Table 3 and the formula will almost certainly be explored by graph plotting and noticing that the points are scattered around a straight line. Here again there are differences in the conventions adopted by physicists and mathematicians of different generations about which all should be aware.

6. Graphical representation

6.1 In paragraph 3.3 above we remarked that teachers of physics and teachers of maths think differently about the relationship $T = -kx$ and the difference is reflected in the way in which their graphs are labelled. The physicist will insist on the need for the units to be specified and this may be done either by writing, for example, "x in metres" or x/metre. In each case he is trying to achieve what he believes his mathematical colleague would wish, that is, that the graph is a relationship between pure numbers. The *quantity* 'x' to the physicist means the pure number x multiplied by a chosen unit length; hence if we divide x by that unit and write x/metre we have reduced the quantity to a pure number.

7. Concluding remarks

The preceding discussion about imagined interactions between teacher and taught in the course of discussions of Hooke's Law might well be characterised as "Much ado about nothing". But to do so would be to brush aside the patient endeavours of more than a decade of effort to raise the standards of teaching in both physics and mathematics. It is not part of our brief to justify or explain the standpoint adopted, but these variations do exist. We hope that this approach will perhaps at least form a basis for discussion between teachers which will help to resolve some of the problems facing pupils as they move between the maths room and the physics lab. There is no question of apportioning truth; the aim is to establish respect and understanding.

Part 2

Acceleration and the idea of a quadratic function

A Physicist and a Mathematician take a close look at an elementary topic.

Charles Taylor
Robert Morris

1. Introduction

1.1 In the discussion of Hooke's Law we considered the varying view points held by Mathematicians and Physicists of different vintages on a simple linear function arising out of a simple physical experiment. It was also recognized that this function was selected as a special case among a wide variety of other functions that could arise in similar experiments. The purpose of this section is to extend the same ideas to a different experiment in which the function turns out to be quadratic.

1.2 In the discussion of dynamics, physicists start from one of Newton's Law of motion and introduce the notion of a force as something that causes a mass to accelerate.

The usual statement of the law is:-

> A body continues in a state of rest, or of uniform motion, unless acted upon by an unbalanced force.

For our present purposes it will be convenient to restrict the motion considered to that along a straight line. Now uniform motion means motion at constant velocity and, if we allow zero as one value of this constant velocity, the part about the "state of rest" can also be omitted. Under these suppositions the law can be re-written:-

> A force is that which causes a body in uniform motion to change its velocity.

1.3 The purpose of our present exercise is first to explore quantitatively the relationship between the change in velocity and the force causing it and then to go on to examine the relationship between the distance travelled and the time, when a body is acted on by a force.

2. The Experiment

2.1 We know that an unsupported object will begin to fall from rest and it follows, therefore, that it is being acted on by a force. We can loosely call this a gravitational force and it is reasonable to assume that the force on each of a series of identical objects is the same.

2.2 We can set up a "ticker-tape" experiment with a trolley drawn by a string passing over a pulley to a very light pan on which we can place one or more of a number of identical objects [e.g. 1 kg weights]. The table can be tilted slightly to compensate for friction so that, if the trolley is pushed gently when there are no weights in the pan, it moves along at constant velocity (Fig. 1(a)).

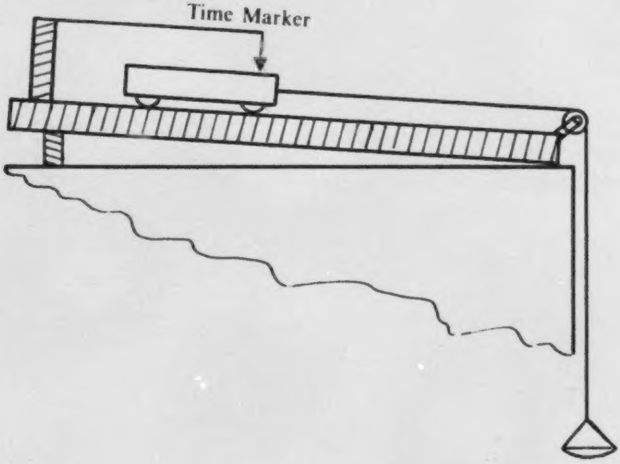


Fig. 1(a)

A time-marker makes marks on the tape fastened to the trolley at equal increments of time. We do not, at this stage, need to know what the increments are, as long as they are equal. And similarly we do not necessarily need to know the size of the weights as long as they are identical.

2.3 On the basis of our discussion so far, it is reasonable to suppose that, if we put 1, 2, 3, 4 etc. of our identical objects in the pan, then the *force* exerted on the trolley will increase in a similar way from 1 to 2, 3 or 4 units.

We now release the trolley from rest with *one* weight in the pan and the tape will look as in Fig. 1(b)

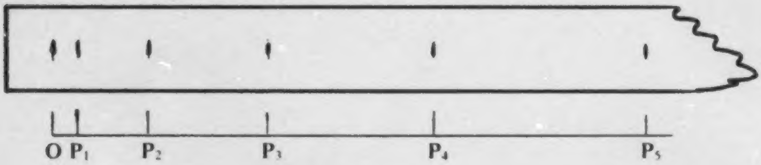


Fig. 1(b)

O is the mark when the trolley was at rest and P_1, P_2 etc. are the successive marks after 1, 2 etc. units of time. Now we measure OP_1 and call this distance a_1 and use *this* as the unit to measure OP_2, OP_3, OP_4 etc.

We shall find that the table of results will be something like this:-

t	measured in 'time marker' units	0	1	2	3	4	5
d	measured in units of a_1	0	1	4	9	16	25

Table 1

The physicist would say that d is obviously a function of t^2 and might write this result in the form

$d \propto t^2$ [read "d is proportional to t^2]

$d = f(t)$ [read "d is a function of t"]

in this case, the quadratic function t^2

A modern mathematician might say that the set of numbers in the upper row map on to the set in the lower row according to the quadratic function $x \rightarrow x^2$

2.4 We now repeat the experiment with two objects in the pan and, provided that we again measure the distances in units of OP_1 which will now have a new value a_2 - the table will be much the same. And so we can successively increase the weight in the pan and produce further similar tables.

We can now prepare a table in which the Force F, measured in arbitrary units is plotted against the values of the initial distance OP_1 in each case.

Force in arbitrary units	1	2	3	4	5
OP (say in mm.)	a_1	a_2	a_3	a_4	a_5

Table 2

and Mathematics Teachers

ICSU Committee on the Teaching of Science

Edited by Alan Rogerson

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED215867

①

Functions and Physics

Robert Morris
Charles Taylor
H. Nagerl
J. Provost
P. Bennett

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

SUSANNE ADLUNG

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SE037055

∫ydx
-b±√b²-4ac
2a
010
001
>
%
×
C₂H₅OH
E=mc²
NaOH

We now measure each of the qualities a_2, a_3, a_4 in terms of a_1 as a unit of length and we find:-

Force in arbitrary units	1	2	3	4	5
OP in units of a_1	1	2	3	4	5

Table 3

In other words, physically speaking, the distance OP_1 is proportional to the force.

Which we can write as

$$a \propto F$$
$$\text{or } a = f(F)$$
$$\text{or } x \rightarrow x$$

and the relationship between the sets representing the force and those representing the distance travelled in the first unit of time is a *linear* function.

3. Complications

3.1 We now need to ask what is the *formula* relating time and distance. We found that $d \propto t^2$ and since d and t are both zero at the same time we can write this as $d = kt^2$ where k is some constant. Now when d and t are both 1, the distance travelled is, in fact, a, and so we can write $a = k$ and hence $d = at^2$.

But suppose, as in fact is more likely, that we cannot be sure that the first mark occurred when t and d were both equal to zero?

We could still extract a meaningful result by using one of the physicist's "tricks", but only if we had a shrewd suspicion in advance that the relationship was going to be quadratic. Suppose we made the measurements corresponding to Table 1 and they came out like this:-

t measured in 'time marker' units	n	n+1	n+2	n+3	n+4
d measured in mm from mark n	0	20	48	84	128

Table 4

How can we treat these results?

One possible trick is based on the fact that

$$n^2 = 1 + 3 + 5 + \dots\dots\dots\dots + (2n-1)$$
$$\text{and } (n+1)^2 = 1 + 3 + 5 + \dots\dots\dots\dots + (2n-1) + (2n+1)$$
$$\text{so } (n+1)^2 - n^2 = 2n + 1$$

We hope that the successive values of d are going to be of the form n^2, $(n+1)^2$, $(n+2)^2$ etc. Hence if we write down the successive differences we should both be able to confirm our hope and also find a value for n. (Table 5).

d measured in mm from mark n	0		20		48		84		128
differences in mm		20		28		30		44	
		=4×5		=4×7		=4×9		=4×11	

Table 5

which confirms our suspicion of the quadratic form and tells us that (2n + 1) for the first pair is 5 and hence that our mark labelled n is in fact the *second* mark after the mark for t = 0. We can now extrapolate backwards to find the distances corresponding to 0, 1 and 2 as in Table 6 and hence we find that the true measure of d during the first units of time (= OP$_1$ = a$_1$) must have been 4 mm and we can then go on and verify the function as before.

Original time in marker units					n		n+1		n+2		n+3		n+4
True time in marker units	0		1		2		3		4		5		6
d in mm	-16		-12		0		20		48		84		128
Differences		4		12		20		28		36		44	
		=4×1		=4×3		=4×5		=4×7		=4×9		=4×11	
d in units of OP$_1$	0		1		4		9		16		25		36

Table 6

The question of units is, of course, an important one, but no new principles are introduced here as compared with the discussion on Hooke's Law.

Part 3

The Mathematical Concept of a Function and the Physical Concept of a Resistance

H. Nagerl
J. Provost
P. Bennett

Introduction

The teaching of Physics at secondary school level depends on the teaching of mathematics in the mathematics classroom and also on the way mathematics is used in the Physics classroom. An undervaluation of the role played by mathematical concepts can lead to conceptual obstacles in the full understanding of physical situations. The importance of mathematical concepts in Physics instruction can be illustrated by the study of resistors.

Discussions with students have shown that they have difficulty in distinguishing between the concept "resistance" and the concept "Ohm's Law".

The resistance R of a conductor is defined as the ratio of the potential difference V across it to the current I flowing through it. That is $R = \frac{V}{I}$. If the potential difference across a metallic conductor is varied and the corresponding current I measured, the ratio $\frac{V}{I}$ is found to be constant so long as physical conditions do not alter. Thus the resistance of such conductors is the same whatever the potential difference applied and we can write $\frac{V}{I} =$ a constant.

This is Ohm's law and conductors, such as metals and alloys, which obey it are called Ohmic conductors. A graph of V against I for them is a straight line.

Many conductors do not obey Ohm's law, that is they are non-Ohmic and their resistance depends on the potential difference. The revolutionary advances in modern electronics are due largely to non-ohmic devices such as transistors. Ohm's law is also not obeyed by some crystals such as silicon carbide, nor by gases as in a neon lamp.

The usual numerical calculations (mainly proportionality calculations) which are valid for the study of Ohmic resistors are not adapted by students to the study of non-linear resistors for which one should introduce calculations on functions (or equivalently, operations on graphs). Without the introduction of the concept of function in the study of resistors, mathematics has been under-used and this leads to conceptual difficulties which could be avoided.

A structured test given to 184 medical students illustrates the confusion between a resistance and Ohm's Law. A series of 5 tasks was given to the students.

Case Study Report

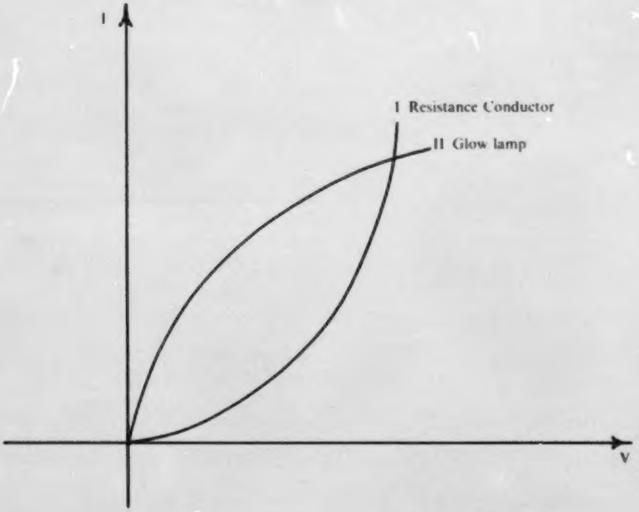


Fig. 1

Task 1. How big is the current through the resistance conductor when one applies a potential difference of 40 V across it? This task was introduced to make sure that students were able to read graphs. 178 out of the 184 students were able to indicate the correct value of the current for 40 V.

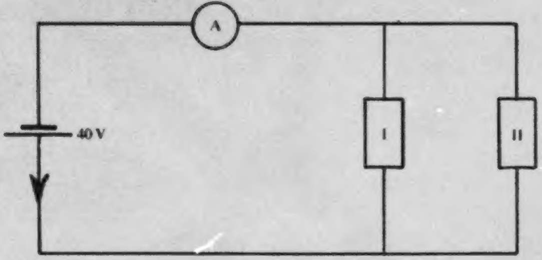


Fig. 2

Task 2. The glow lamp (characteristic II) and the resistance conductor (characteristic I) are connected in parallel. Which value of the current is indicated by the ammeter?
(The p.d. across each resistor is 40 V).

This task as well as the next one is concerned with the additivity of currents (Kirchoff's Law $\Sigma I = 0$ at a node). From the 178 students correctly solving Task I, 12 did not attempt this task nor the following ones. 68 students presented the expected solving method: they read the value of current at 40 V for each characteristic and added these two values:

$$\begin{aligned} I &= I_1 + I_2 \\ &= 195\,\text{mA} + 162\,\text{mA} \\ &= 357\,\text{mA} \end{aligned}$$

The 91 remaining students used an awkward but instructive method. They read from each curve the values of the current and calculated the conductances ($L = \frac{I}{V}$), for example $L_1 = \frac{195\,\text{mA}}{40\,\text{V}}$. They then made use of the formula for resistors in parallel $L = L_1 + L_2$ and obtained the current by multiplying the calculated value of L by the voltage 40 V. These students will be referred to as "calculators". They give a key role to the notion of resistance (or conductance) as soon as *operations* on resistors are involved. They obtain the current, voltage or conductance from the unique formula (between numbers) $I = LV$. In the form $L = \frac{I}{V}$ this formula is used to get the conductance from the reading of the characteristic. The fact that L depends on V is not clearly realised. Therefore in Task 3, many of them think of the relation $I = LV$ in terms of a linear relation between I and V.

Task 3. Draw the current-voltage characteristic of the parallel connection (Fig. 2) on Fig. 1.

This task can be solved with the same strategy as before. However it involved new problems for the students.

From the 68 students who solved Task 2 in the expected way, 23 did not attempt this task. The other students (45) answered correctly ("addition" of the two curves).

From the 91 "calculators", 48 did not attempt this task. 35 students draw an Ohmic linear graph associated with the previous calculation of resistance at 40 V (wrong answer). Only 8 students were able to solve the problem, 7 by coming back the expected way and one by the awkward strategy he used on Task 2.

For many of the students operations on resistors are not considered as implying operations on voltage – current curves. This is clear for "calculators" who use these curves only to deduce the values of the conductances, but it is also the case for those who have difficulty in inferring from the relation between numbers $I = I_1 + I_2$ a relation between curves. Some students among those who omitted Task 3 probably wonder whether a current-voltage characteristic can be defined for resistor networks.

Task 4. and Task 5.

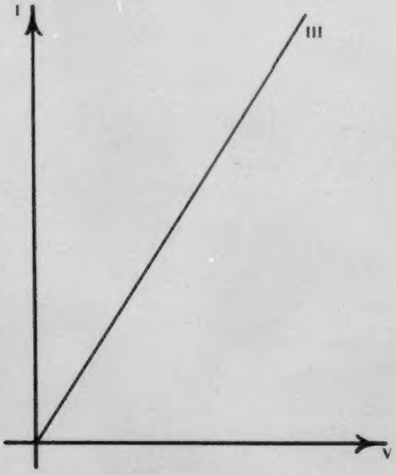


Fig. 3

An Ohmic conductor (Characteristic III, Fig. 3) and the resistance conductor are connected in series (Fig. 4).

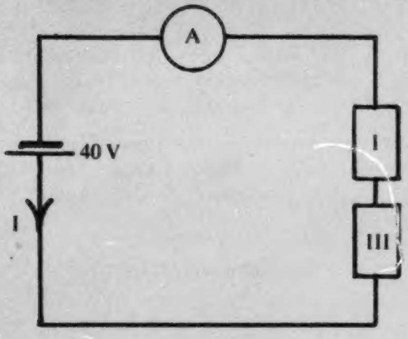


Fig. 4

The potential drop over the two conductors taken together is 40 V. What current flows through both resistors?

The Ohmic resistance of Fig. 4 is replaced by the glow lamp and the potential difference of 40 V is applied to this combination. What is the current flowing through both resistances?

A hint is given to possible method of solution of these tasks by suggesting interchanging the abscissa and ordinate of the characteristic graphs, i.e. plotting V against I.

The groups of 91 "calculators" failed with these tasks completely: they either omitted the tasks or calculated the resistances as before. From the group of 68 students, only 2 gave the right answer: they computed graphically the characteristic curves of the series circuits as suggested by the hint in Task 5. 8 students read the value of current from only one curve (wrong answer). The other students either gave no answers or came back to calculation methods used by "calculators".

Although the hint accompanying Task 5 suggested a graphical solution, the difficulty of these tasks has increased the percentage of "calculators". Even the majority of the 68 students who solved Task 2 in the expected way came back to calculation methods indicating that they do not completely comprehend operations on voltage–current curves.

The systematic use of conductance by calculators should be traced back to the idea that physical laws deal with physical quantities and that calculation on these quantities essentially involves numbers. A resistor therefore should have a resistance (number) and the resistance of an association of resistors should be obtained from some combination of the resistances of individual resistors. The existence of a relation between current and voltage is understood for practical tasks as a collection of values of conductances. The resistor is not conceptually considered by students as some (materialised) function which gives current values in terms of voltage values (or maps current values and voltage values). This explains why students who were able to read graphs in Task 1 and write $I = I_1 + I_2$ in Task 2 failed in constructing the characteristic curve of the parallel circuit: Task 3 implied an operation (addition) on functions (not only addition of individual numbers).

From the characteristic curve of a resistor, two functions can be defined. One function gives the potential difference according to the values of the current. Calling this function r, we have $r : I \to V$ so that $V = r(I)$.

This function can be called the "resistance function".

The set of potential differences is called the *domain* and the set of current values the *range*.

Ohm's law can be stated as r is the function "multiply by R"; $r : I \to RI$.

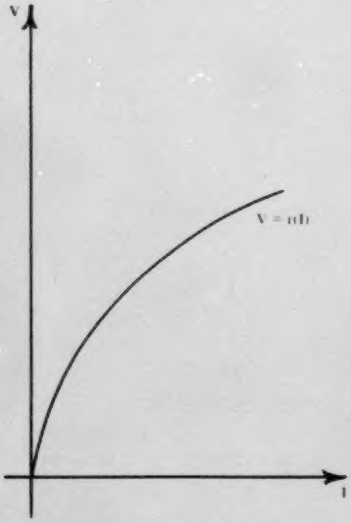


Fig. 5

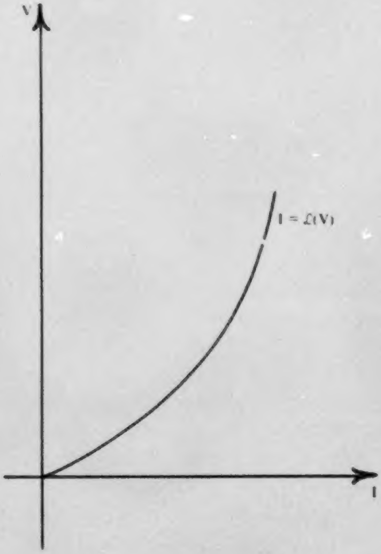


Fig. 6

The graph of r is shown in Fig. 5 for a non-Ohmic conductor. In a similar way the "conductance function" $\mathcal{L}$ can be defined $\mathcal{L} : L \to I$ so that $I = \mathcal{L}(V)$.

The graph of $\mathcal{L}$ shown in Fig. 6 is the characteristic of the resistor.

For an Ohmic conductor $\mathcal{L}$ is the function "divide by R"; $\mathcal{L} : V \to \frac{V}{R}$. Both $\mathcal{L}$ and r are one-to-one mappings (i.e. each element of the co-domain is associated with only one element of the domain and vice versa). r^{-1} and $\mathcal{L}^{-1}$ therefore exist.

$$r : I \to V \text{ implies } r^{-1} : V \to I$$

The functions $\mathcal{L}$ and r satisfy the functional relation

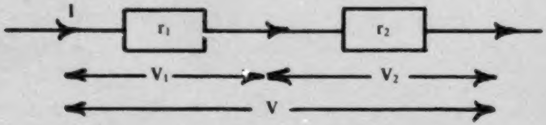
$$\mathcal{L} = r^{-1}$$

(or $\mathcal{L}$ is the identity function $\mathcal{L}r(I) = I$).

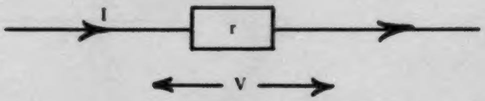
The curves representing the two functions $\mathcal{L}$ and r on a graph are reflections of each other in the line at 45° to the V axis.

A resistor is a black box defined by the function r (or $\mathcal{L}$). It is an easy task to find the resistance function of two resistors in series and in parallel.

In series



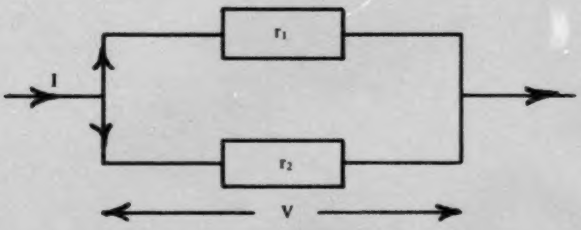
We wish to replace the combination of the two resistors by a single resistor r



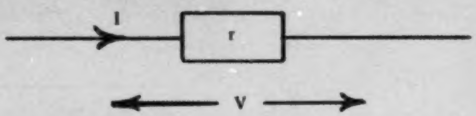
$$V_1 = r_1(I)$$
$$V_2 = r_2(I)$$
$$V = r(I)$$
$$r(I) = V = V_1 + V_2 = r_1(I) + r_2(I) = (r_1 + r_2)(I)$$

so that we have equality between *functions* $\underline{r = r_1 + r_2}$

Resistors in parallel



Replacing the combination of two resistors by a single resistor r



$I_1 = \mathcal{L}_1 (V)$ where $\mathcal{L}_1 = r_1^{-1}$
$I_2 = \mathcal{L}_2 (V)$,, $\mathcal{L}_2 = r_2^{-1}$
$I = \mathcal{L} (V)$,, $\mathcal{L} = r^{-1}$
By $I = I_1 + I_2$
$\mathcal{L}(V) = \mathcal{L}_1(V) + \mathcal{L}_2(V) = (\mathcal{L}_1 + \mathcal{L}_2)(V)$

so that we obtain the equality between functions $\underline{\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2}$

or in terms of the resistance functions $\underline{r^{-1} = r_1^{-1} + r_2^{-1}}$

The above formulae have a striking similarity with the usual laws for Ohmic resistors and therefore need little effort of memory. They allow us to make graphic operations on the curves associated with the resistance and conductance functions.

Of course, if the function r means "multiply by R" (R a fixed number), we recover the usual formulae for Ohmic conductors in series and parallel

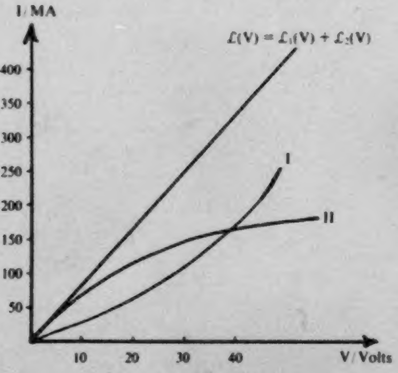
$$R = R_1 + R_2$$
$$L = L_1 + L_2$$

These formulae of course involve numbers and not functions.

Task 2 - 3
For 2 resistors in parallel we have

$\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$
$\mathcal{L} : V \to I$
$\mathcal{L}(V) = (\mathcal{L}_1 + \mathcal{L}_2)(V)$
$= \mathcal{L}_1(V) + \mathcal{L}_2(V)$

So to get the characteristic for the combination we have to "add" the two characteristics

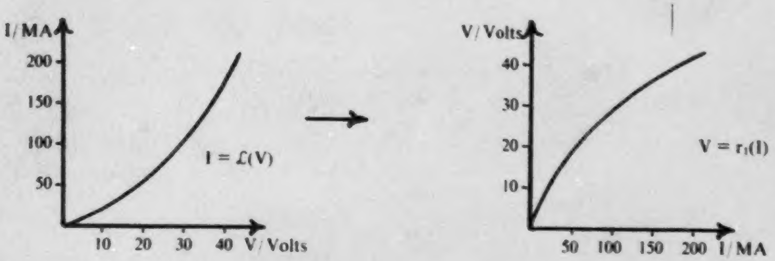


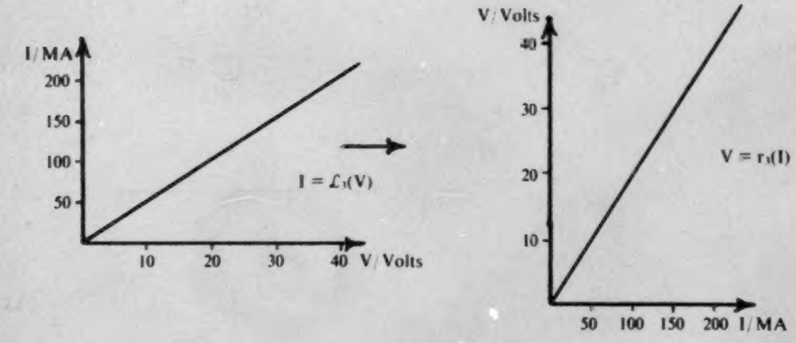
Task 4 – 5

For resistors I and III in series we have

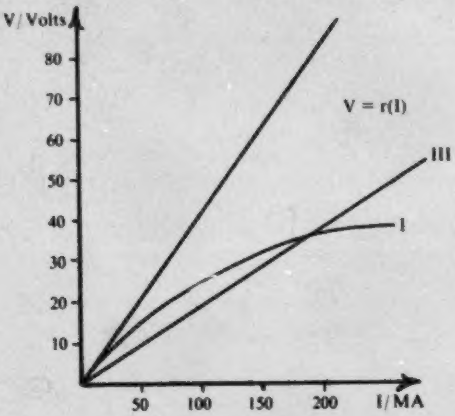
$$
\begin{aligned}
r &= r_1 + r_3 \\
r &: I \to V \\
r(I) &= (r_1 + r_3)(I) \\
&= r_1(I) + r_3(I)
\end{aligned}
$$

r is the inverse function of L. The graphs of $r_1 = I \to V$ $r_2 : I \to V$ can be obtained by reflecting the characteristic curves in the line at 45° to the V axis.

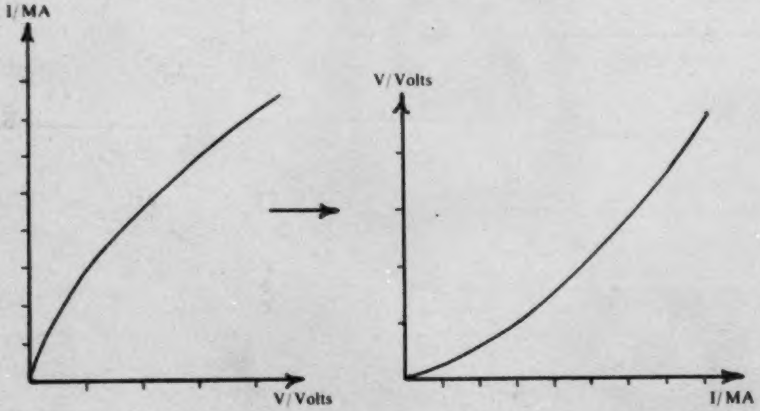




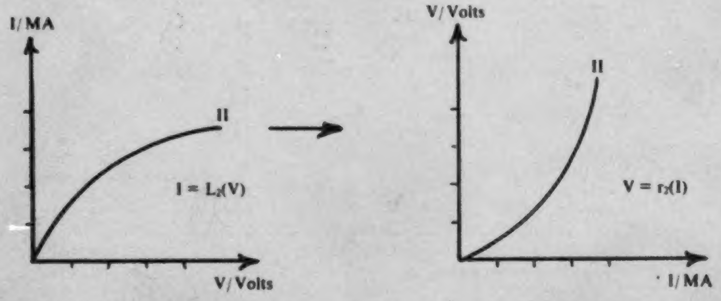
r is obtained by adding r_1 and r_3.

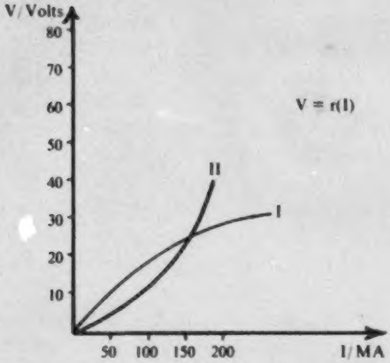


The characteristic is then obtained by reflecting in the line at 45° to the I-axis.

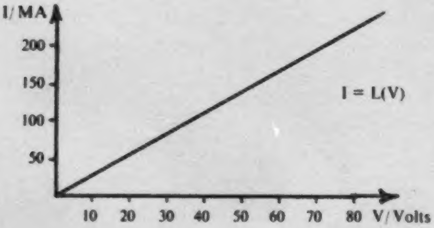


Similarly for resistors I and II in series





Again the characteristic can be obtained by reflecting in the line at 45° to the I-axis.



The usual presentation of Ohm's law in secondary school texts has probably caused many of the students' problems. This presentation does not clearly exhibit that Ohm's law is a very particular law. This can be clearly exhibited only in a functional approach since it is well known that linear functions are very special functions. This law written in the standard form V = RI is unconsciously considered by students as the numerical version of the quantity calculus formula [V] = [R][I]. Since this last formula has a general validity, it needs a small mental jump to think that Ohm's Law is also a general law.

The physical concept of "linear law" although trivial in appearance cannot be comprehended or taught without the mathematical concept of a function. Confusion is often encountered between physical concept and that of linearity (additivity) of physical quantities. For example, students easily think (and books generally help them to do so) that the additive (linear) character of the quantities "current" and "voltage" exhibited by Kirchoff's laws and of a general validity, implies the existence of a linear law (Ohm's law) between these quantities. This is the assumption made by the 35 students who drew an Ohmic linear graph in Task 3. In the same way students are reluctant to add non-linear functions as shown by the 23 students who were able to add values of current in Task 2 and were unable to extend this to "add" the graphs in Task 3.

Given two functions f_1 and f_2 their sum is defined as $f_1 + f_2 : x \rightarrow f_1(x) + f_2(x)$ even if the functions are non-linear ones.

The transfer of knowledge of Ohmic resistors to the study of non-Ohmics (i.e. non-linear) resistors can be greatly aided by the introduction of the concept of a function.

The introduction of the mathematical concept of function in the Physics course is not motivated by the will to use "modern" Mathematics as a universal language for Physics. It is not the word "function" which is important in itself but those operations which are possible on functions and which constitute the best model for those operations which are done on physical resistors. If "modern" Mathematics is useful in Physics, and we think it is, this should be experienced by Physicists (as for example in the test previously described). These experiments will in turn make the teaching of mathematical concepts a more concrete one for pupils.

Acknowledgment

The authors are grateful to the Deutsche Forschungsgemeinschaft for support.

This series of booklets arose out of a conference held at Bielefeld in 1978 and which brought together teachers of Mathematics and of the various Sciences at a variety of levels.

The Committee on the Teaching of Science of the International Council of Scientific Unions (ICSU-CTS) has published the series and is grateful to all the contributors, to the Editor and to those who sponsored and organised the Conference. Special thanks are due to UNESCO for the financial assistance that has made publication possible.

① **Functions and Physics**
② **Links Between Geography and Mathematics**
③ **Our Inheritance: Common Ground for the Mathematics and Biology Teacher**
④ **Mathematics and Chemistry**
⑤ **Mathematical Modelling**
⑥ **Mathematical Modelling with Calculus**

Printed in England by John Goodman & Sons (Printers) Ltd., Cardigan Street, Birmingham B4 7RS

ICSU Committee on the Teaching of Science

Edited by Alan Rogerson

②

Links Between Geography and Mathematics

Gerard Larkin

∫ydx
010
001
-b±√b²-4ac
2a
≥
×
÷
C₂H₅OH
E=mc²
NaOH
28

This positive and stimulating booklet will be of interest to both the mathematician and geographer. It presents three detailed case studies in the form of pupils' projects which clearly illustrate how mathematics and geography can be effectively linked at school level. It then supplies a detailed list of many situations in geography that are available for quantitative treatment, with the corresponding mathematics links.

The booklet provides a practical incentive to the geography teacher to illuminate his subject using mathematical models, and to the mathematics teacher it provides a range of real applications for his subject from geography.

Links between Geography and Mathematics

Gerard Larkin

Head of Geography
St. Maur's Convent
Weybridge

Editor's Preface

What is "new" about the 'new geography? Instead of being largely descriptive in method with the *region* as its chief object of study, it attempts to use and apply mathematical terminology, techniques and models to social and economic factors rather than the historical or physical. The methodology of the new geography is analytic, using concept formation, modelling, and data collection and representation as its main tools. During the past fifteen years the revolution in geography teaching at university level has had an increasing effect on schools. The 1974 DES education survey, *School geography in the changing curriculum*, reported that about half of the 217 schools sampled were using a new syllabus. A smaller percentage of schools used newer styles of geography teaching (e.g. games and simulations), but this proportion seemed likely to increase during the next few years.

The new geography owes much of its impetus to a series of seminars (1963-66) held at Madingley Hall, Cambridge. Chorley and Haggett's *Models in geography* (Methuen, 1967) brought much of the current research to the notice of a wider audience at both school and university level. Cole and Beynon's brilliant *New ways in geography* (Blackwell, 1968-   ) showed how these new ideas might be introduced into schools, and they blazed the trail which more writers are now following. The most interesting of these are the Walker and Wilson series *Location and links* (Blackwell, 1973), an excellent introduction to the new geography in *Science in geography* (Oxford University Press, 1974), Dalton's *Simulation games in geography* (Macmillan, 1972) and Walford's *Games in geography* (Longman, 1969).

The new geography has obvious significance for the mathematician, and many of the mathematical topics it uses are emphasised in 'modern' mathematics at school and university: functions, transformations, probability and statistics, networks and numerical methods. The most exciting aspect of the new geography is the wealth of practical situations in which mathematics is relevant and useful: surveys and questionnaires to determine the siting of new supermarkets or factories, the concept of 'distance' using different measures, topological ideas and networks for designing urban traffic systems or discussing human settlements. The development and application of mathematics in geography has reached a stage (passed in economics and still to come with biology) where initial communication and collaboration between geographers and mathematicians would be mutually beneficial. Mathematicians gain a fertile field of real-life situations to illustrate mathematical techniques, while there are substantial benefits for geographers in relating the applicable mathematics to their subject in an intelligible way.

Alan Rogerson

Links between Geography and Mathematics

As in Mathematics, so in Geography — the key ideas of the subject recur several times in the course of a Secondary School course of instruction, each time at a higher level of complexity — the 'spiral view' of the curriculum.

Many of the concepts in Geography are best expressed in terms of Mathematics — numerical models, geometric patterns, statistical techniques of handling a huge and varied mass of data, etc. Geography has recently acquired a body of theory which expresses many of its basic concepts in mathematical terms, but this has penetrated first at Sixth Form Level, necessitating a change in method of thinking from literary/description/illustrative to analytical/mathematical between the Fifth and Sixth Form levels.

This booklet hopes to indicate that simplified (but still valid) versions of the Mathematical approach can be introduced in the 11-16 age range, forming the groundwork for later elaboration, providing that they are mathematical models which *really illustrate and clarify the geographical concepts*, rather than being a froth of 'quantitative techniques' adapted willy nilly in confused ways to give the subject a spurious air of scientific authenticity.

The booklet consists of:

A. One or two examples indicating in more detail the kind of work envisaged from the children's point of view. Two typical cases are dealt with. The 'Christaller landscape' is a division of a surface around points, elaborated originally as an analysis of the market areas of the hierarchy of settlement in Germany. It contains solid principles of subdivision of an area which can be applied to other fields of Geography. It is, at present, sprung on pupils at Sixth Form level: but it could be built into their thinking at lower levels, third or fourth form, for example, by careful logical development in simple terms.

   The 'Water Cycle' is taught as a system to pupils in Biology and Geography but almost purely in diagrammatical/pictorial form. It is possible to state it also in quantitative form, and to introduce simple equations with theoretical large-scale and practical small-scale examples, leaving pupils with a much clearer view of the workings of the system. The Hydrology examples attempt to do this.

B. An outline of some syllabus areas, with indications of where mathematical expression may be applicable (i.e. areas worthy of further investigation).

In the ideal case there is as much a need for a 'Geography for Mathematicians' approach as for a 'Maths for Geographers' approach. The co-operation of Geographers and Mathematicians in a single school could take the form of a joint approach to topics such as the $Q = P - E \pm \Delta S$ formulae. (See Hydrology below), the Geographers providing the appreciation of the physical framework, the mathematicians the manipulative techniques. Both would benefit from the co-operation.

FOOD AND SETTLEMENT

General

Food is produced on the land, and consumed (generally) in people's houses. At its simplest, the food produced over an *area* must travel along a number of *lines* to reach a *point*. Geography can be seen as the interaction of areas, lines and points. Look at Fig. 1: here is a field of wheat — an area — the solid line shows the harvested wheat being collected by a truck, and it must all converge on the gate — a point.

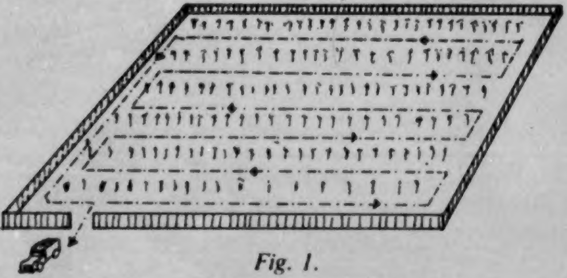


Fig. 1.

Clearly the reverse movement is also applicable: wheat has converged on the flour milling factory, and the flour must be distributed to the bakers.

This happens on many different scales but is a basic principle. Where people live depends on food and transport — can food be brought to the people, or must people live where the food is to be obtained?

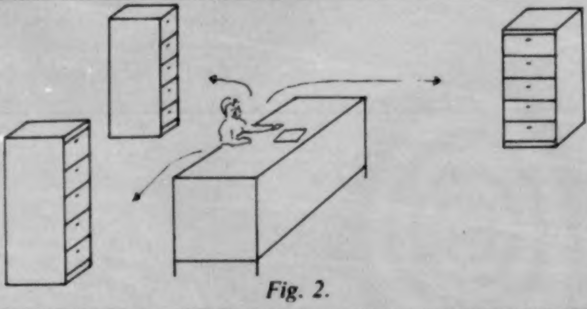


Fig. 2.

Look at Fig. 2. It shows a man working in an old-fashioned office. He needs to get papers from large filing cabinets, which are heavy, and rest on a rough floor. He cannot move the cabinets to his desk, so he must go to the cabinets to get the relevant set of papers. In a modern office the cabinets move on well-oiled castors across a smooth floor, so he can move the relevant cabinet to the desk and work on the papers.

In early days in our society (and still in many Third World countries) transport was difficult and inefficient — things had to be carried on foot, or on horseback slowly, and along bad roads. Most food goes bad easily and the means of preserving it were not known. We can consider distance as a kind of *friction* which makes movement difficult — like the cabinets on the rough floor. More efficient transport is like putting castors under the cabinets and smoothing the floor. We can imagine it as *lubricating* the system to eliminate the friction. The more this happens the more it becomes possible for food and other material to be brought *to* the people, and the more they are able to concentrate in large cities. (Whether they want to, or whether we think they should, are different questions: they are *able* to). So a great deal of historical and geographical change can be explained as the result of lubricating the system to eliminate the 'friction of distance'.

Case 1. Nomads

Consider the Fulani people who live on the savannas of West Africa. They live in areas where the rainfall is insufficient, and too erratic to allow permanent farming of the land. They graze cattle on poor grassland; the animals quickly eat all the good grass in any area where they graze, and the Fulani herdsman can only keep them alive by moving them to new areas of,pasture. When they are moving with their cattle they live in tents and their dwellings can be thought of as a number of *points along a line*. The line is likely to be sinuous ('snaky') to use up all the area, and the area grazed by the cattle can be imagined as a large number of circles along the line, (See Fig. 3).

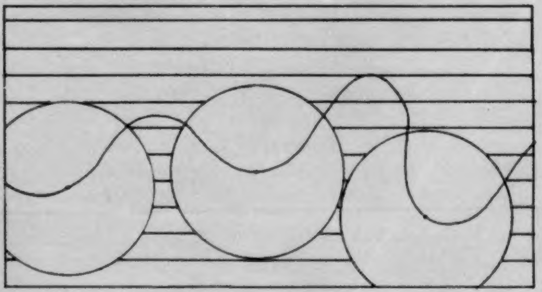


Fig. 3.

Fig. 3 might also show the path taken by a travelling circus. Each dot is a town where the circus stops for three days. The surface of the paper is the area, which in this case is all the populated land where there are potential customers for the circus. No one wants to see the same circus every week, or to travel a very long distance to see it, so within (say) 10 km. of the circus there will only be a limited number of customers.

> *Exercise*: suppose there are 1000 people living in every square kilometre of the area; how many people will live within a radius of 10 kilometres of the circus?

Again because of the 'sparseness' of the customers, the 'point' must move for the circus to exploit it efficiently. If the circus were to stay permanently in one town it would quickly use up its potential customers (like overgrazing the savanna!); so the circus performers are also nomads.

Case 2. Subsistence Farmers

Suppose we look at an area of peasant subsistence farmers, in an area where families work with simple tools to get the necessities of life for themselves (food, clothing, timber for building and fuel) with little or no sale of goods or trade for profit. We wish to find out:-

1. What shape will the farms be?
2. What size will the farms be?
3. Where will the farmer put his house within the farmland?

In many cases, physical geography will decide this. Scottish crofters located their farms in the sheltered valleys, or along raised beaches by the side of the lochs. But suppose there are no important physical barriers to settlement and farming, that we are looking at a flat plain of even fertility where transport is equally easy (or difficult) in all directions? Here mathematical rules linking area to points and to lines will help us to understand.

Question 1.
Where do the people live? Clustered in one area, or spread over the land on their farm?

Since we are supposing a society with very limited technology which must apply to transport, there is a high degree of 'friction of distance': in other words it is impossible for food to be rapidly and efficiently assembled at one point from large areas, and it is impossible for workers to live far from their fields. So the market for the produce of the area (i.e. where the food, fibres and fuel are used) will be a large number of points, and movement will be a very large number of small-scale movements inward, like the radii of circles — radial movements. (See Fig. 4).

Thus the farms must be located among the fields.

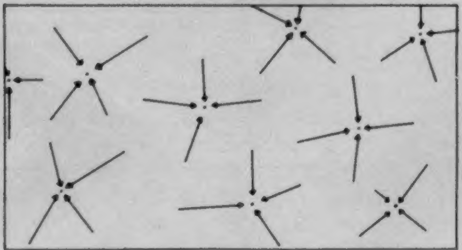


Fig. 4.

Question 2
Where will the farmhouse be located within the farm?

One of the farmer's biggest problems is the time spent getting to and from the fields each day with his tools and farm implements.

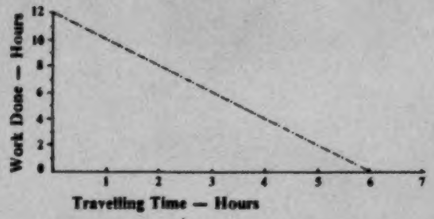


Fig. 5.

In Fig. 5 the *y* axis of the graph shows the amount of productive work done for a 12 hour day, the *x* axis the number of hours walking time (one way). The field next to the farm at 0 produces 12 hours *useful* work for a 12 hour day, while a field at six hours walking distance away cannot be farmed because no useful work is done. We can draw two important conclusions from this:-

1 It is very much in the interests of the farmer to have all his fields as near as possible to his farmhouse. (Countries like France which have a farmer's land broken up into many small scattered strips are now in the process of concentrating the land in compact blocks).

2 Land which the farmer may acquire on the fringes of his farm is of less and less use to him the further away it is. This is one example of the law of diminishing returns.

Question 3 What shape is most efficient?

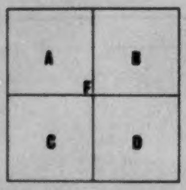


Fig. 6.

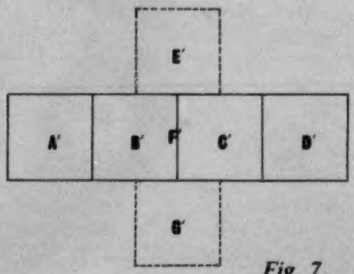


Fig. 7.

Consider the two possible farm shapes in Fig. 6 and Fig. 7. Each shows a farm of 4 fields; each field is a square 100 metres by 100 (a hectare).

Exercise: calculate the distance from the farmhouse F (F^1) to the furthest point on the farm for Fig. 6 and Fig. 7.

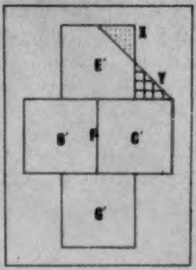


Fig. 8.

Supposing the farmer had a chance to exchange two fields E and G in Fig. 7 for his fields A^1 and B^1, and that they are of equal size and fertility. Will he make the exchange? Why?

Now suppose that in Fig. 8 he had the chance to exchange area Y for his area X — again the two being equal. Will he exchange? What shape *does* have all its area within the shortest possible distance of a centrally placed farmhouse?

The answer is, of course, a circle.

> *Exercise*: calculate the radius of a circle which would contain the same amount of land as the farm in Figs. 6 and 7 and 8.

Question 4: what will be the size of the farms?
We are assuming that:- a) Each family works its own land with simple implements and has no mechanised transport.
b) No employed labour is available.

There will be a *minimum* size determined by the amount of produce needed to supply the family with their needs. (NB. the actual size needed by a family varies according to the fertility of the land with the techniques available). There will be a *maximum* size determined by the area of land which the farmer's family is capable of farming with the time and energy at their disposal (taking into account the travelling time, as above). If the farm is circular, both the minimum and maximum areas can be expressed as radii. (Fig. 9)

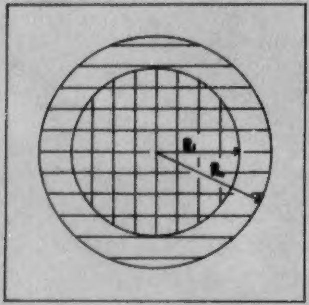


Fig. 9.

Suppose the family *can* farm 5.5 hectares, and needs to farm 3.5. Calculate the radii R_1 (minimum) and R_2 (maximum).

With only simple tools, and no employed labour the maximum and minimum areas will be very close together. If these conditions exist all over the area, then the farms will all be of *approximately the same size*, and the farmhouses in the centre of the farms all *approximately the same distance apart*.

Question 5: What will be the disposition of the farms on the land, what will be the *pattern* they make?

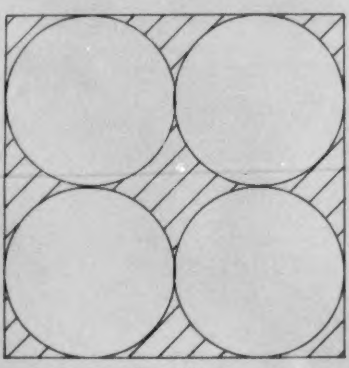


Fig. 10.

Look at Fig. 10. Circular farms cannot fit on an area without leaving unused land (shaded in Fig. 10). If we put the circles on a rectangular grid (Fig. 10) we waste more land than if we put it on a triangular grid (Fig. 11); you can try this by stacking glasses of the same size on a narrow shelf; if you put the second row of glasses exactly opposite the first, you will not fit as many rows on the shelf as if you put the second row in the spaces. Look at the pattern you have made; it is the same triangular grid as Fig. 11.

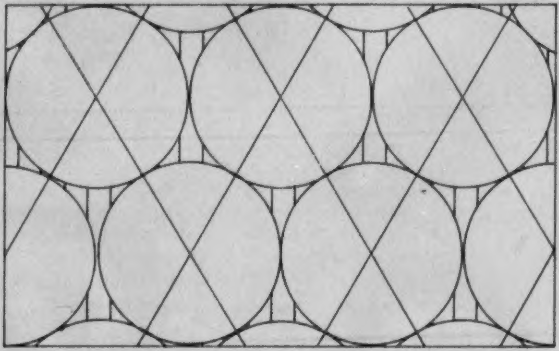


Fig. 11.

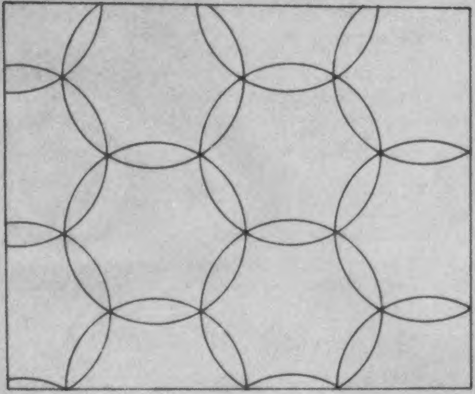


Fig. 12

Suppose we move the circles closer together, as in Fig. 12, we end up with some land shared by two farms. Draw this diagram and show the construction needed to give this shared land to the *nearest* farm. What pattern have you made?

So – we end up with a tesselated pattern of regular hexagons. These are the mathematical rules to answer the question:

> How can the area most efficiently be divided up between a number of points?

This *theoretical* answer is the same whether we are thinking in terms of the disposition of peasant villages in Nigeria, or the setting up of a distribution network for cars in a new market area. Whether it is *point serving area* (selling and servicing cars) or *area serving point* (agricultural production) the principle is the same.

This section does *not* attempt to prove that farm boundaries in peasant areas will be hexagonal, or that the only efficient distribution network will be a tesselated pattern of regular hexagons! It does try to introduce a theoretical element into the understanding of locations. Geography has explained most distributions in terms of transport lines and nodes, and the limitations of physical geography. These explanations become blurred when areas of flat topography with no marked variations in fertility or ease of transport are concerned. So it is perhaps clearer to think in the opposite direction:-

1. **What pattern will emerge purely in terms of area, line and point?**
2. **What variations from this ideal pattern can be seen, and how can these be explained in terms of:-**
    - **a) Physical variations of the landscape.**
    - **b) Transport variations**
    - **c) The forces of agglomeration allied to physically or randomly controlled initial variations?**

TRANSPORT - CENTRALITY

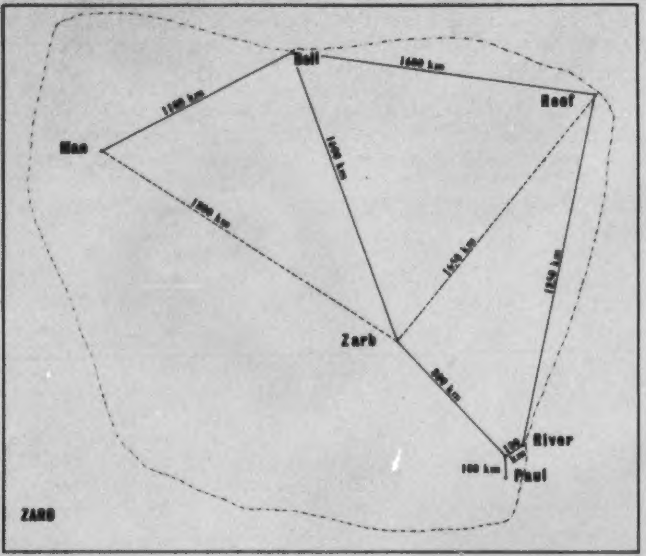

Bell
1600 km
1100 km
Reef
Man
1400 km
1800 km
1650 km
1250 km
Zarb
900 km
100 km
River
100 km
Paul
ZARB

Fig. 13.

Look at the map, Fig. 13. This is the country of Zarb. In this country there are 5 main towns, Man, Bell, Reef, River and Paul. The government of the country wants to set up a tax office so that its chief inspector can visit *every* town on a regular basis. In which town should the Government set up the office? You should be able to answer this question by completing this route matrix by filling in the distance travelled along the roads shown (only those marked by solid lines have been built).

	Man	Bell	Reef	River	Paul	Total
Man	0	1100	2700	3600	3600	11000
Bell	1100	0	1600	2500	2500	7700
Reef	2700	1600	0	1850	2050	8200
River	3600	2500	1850	0	200	8150
Paul	3600	2500	2050	200	0	8350
Total	11000	7700	8200	8150	8350	0

Matrix 1

It should be obvious that this is a very difficult country to govern because of the vast distances involved, and the way in which the towns are spread out along the fringe of the country. The government is considering building a new capital to be called Zarb, nearer the middle of the country. Complete Route Matrix 2 to find out the effect of this.

Another proposal being considered is the building of new roads, shown in dotted lines on the map. Complete matrix 3 and recalculate the distances using the new roads.

Matrix 2

	Man	Bell	Reef	River	Paul	Zarb	Total
Man	0	1100	2700	3600	3600	2700	13700
Bell	1100	0	1600	2500	2500	1600	9300
Reef	2700	1600	0	1850	2050	2750	10950
River	3600	2500	1850	0	200	900	9050
Paul	3600	2500	2050	200	0	900	9250
Zarb	2700	1600	2750	900	900	0	8850
Total	13700	9300	10950	9050	9250	8850	0

Matrix 3

	Man	Bell	Reef	River	Paul	Zarb	Total
Man	0	1100	2700	2700	2700	1800	11000
Bell	1100	0	1600	2500	2500	1600	9300
Reef	2700	1600	0	1850	2050	1650	9850
River	2700	2500	1850	0	200	900	8150
Paul	2700	2500	2050	200	0	900	8350
Zarb	1800	1600	1650	900	900	0	6850
Total	11000	9300	9850	8150	8350	6850	0

Write a brief report to the Government, covering these three points:-

a) which is the best existing town in which to set up the tax office?
b) What would be the effect of building the new capital?
c) What would be the effect of building the new roads?

In each case indicate the *actual* figures for travelling distances which are involved.

We have been calculating the *Centrality* of the towns.

(This example has been taken from an actual country, and the names only lightly disguised: see if you can decide which.)

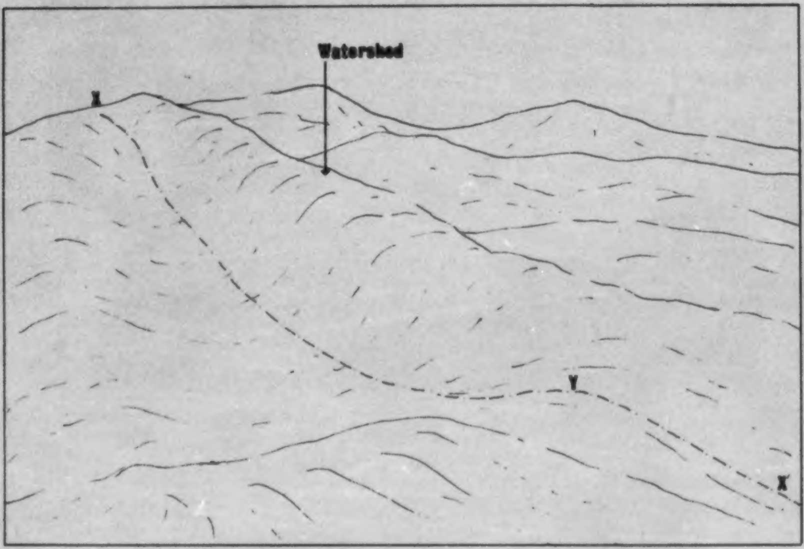


Fig. 14.

Fig. 14 shows an area which has just been pushed up out of the sea. Do you think that new land formed by folding and faulting will have all the land sloping towards the sea to get rid of surplus water? Obviously not; there would be some hollows which do not drain out to the sea. Let us consider such a hollow.

Rain falls on the land. What happens to it?

a) Some sinks into the pore spaces and cracks in the rocks. Obviously these spaces cannot absorb an infinite capacity of water: they have a *storage capacity*. When the spaces are full you cannot pour any more water into the rock, and any surplus will simply run away. If no more rain falls for a long time the water will simply seep out and run away.

b) Some will evaporate, or be transpired by plants back into the atmosphere as a vapour. Since the end result is the same, we combine these as *evapotranspiration.* (e/t)

c) Some will run off the land into the streams and rivers.

Obviously since water does not simply disappear and only under rare conditions can new water be made we should be able to trace where all the water falling in an area ends up. Here is a simple equation to help:-

Q	=	P	—	E	(±) S
Runoff – all the water running *out* of the basin – i.e. the surplus		Precipitation – all the rain snow, sleet etc. falling on the area of the basin		All the water evaporated or transpired back into the atmosphere – strictly e/t.	Changes in storage; rain after dry weather seeps into the spaces in rock and soil, and you have *less* surplus than expected – i.e. the sign is negative; rain after wet weather produces more surplus than expected, and the sign is positive.

In September 1976 after a long drought came a period of moderate rainfall; would you expect the S sign to be negative or positive? In February 1978, after a wet winter came a similar period of moderate rainfall; would you expect the S sign to be negative or positive?

Obviously in the long term, storage will even itself out. So we will ignore it when talking of rainfall over a year, and use the simpler equation $Q = P - E$.

Let us return to the area of land draining into the hollow. Here are two cases of different climates: answer the questions filling in the relevant sections of the table below as you do so.

Fixed data: area of land draining into the hollow = 150 km^2

Q	P	E
Case 1	P = __________	Land Surface
		E = __________
		Plus Lake Surface
		E = __________
		Total
Q = __________		E = __________
Case 2	P = __________	Land Surface
		E = __________
		Plus Lake Surface
		E = __________
		Total
Q __________		E = __________

(All volumes to be given in m^3).

Case 1 Rainfall = 250mm per year over the whole land surface.
E/T = 200mm (the *actual* amount of water lost from the land surface)

Question 1: What is the total volume of water which falls in the basin?
2: What is the total volume of water which is evaporated?
3: How much surplus water is available?

This surplus water will collect in the bottom of the basin to form a lake.

Fixed data: the greatest possible area of the lake surface before it overflows = 9 km^2.

From the *water* surface 1250mm of water is evaporated in a year (much more than from the land because the water is always *there* to be evaporated). As the lake rises its surface area increases, and so does the total amount of evaporation.

Question 4: Calculate the maximum size of the lake under these conditions. (Don't forget that you cannot have evaporation from land which is covered by a lake!)

5: Calculate Q, the surplus water from this basin under these conditions.

Case 2 Suppose that the climate gets damper and cooler over the years. As the rainfall increases evapotranspiration decreases. (You should be able to find at least two reasons why).

Rainfall = 500mm
E/T on land = 100mm
Evaporation over water = 500mm

Question 6: Calculate the figures for P, E and Q, and enter them in the table.

In Case 1 there is no surplus water (Q) and so *no external drainage*. This has two important results:-

A. Salts weathered out of the rocks and carried down by the rivers remain in the lake in the centre of the basin when evaporation removes pure water. So the lake will be a salt lake, and the salt content will increase as time goes on (perhaps to the stage at which some of it has to be precipitated as salt layers around and beneath the lake). In Case 2 salts are carried away in the river, and the lake will only have the tiny amounts present in river water – i.e. it is a freshwater lake.

B. The rest of the rivers load, the solid particles, cannot be taken out of the basin in Case 1 either, and will accumulate in the lowest part of the basin. In desert areas such as this sheets of sand and finer sediments fill up the basins, covering up what is left of the high land, and forming *continental* sedimentary rocks, instead of the more usual *marine* sedimentary rocks formed when the sediments are taken out of the basin in rivers, and dumped in the sea (as in Case 2).

In Case 2 there is a considerable surplus of water which will eventually overflow the basin as a river. Fig 15 is a profile drawn along line X - X¹ shown in Fig. 14. The river is shown overflowing at point Y.

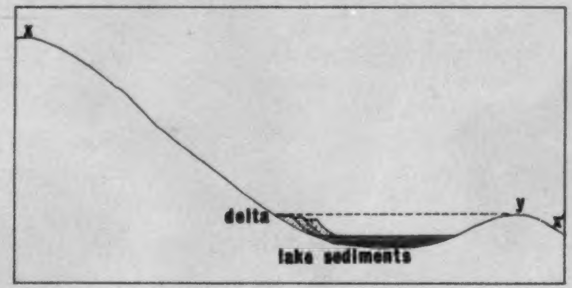


Fig. 15.

Question 7: How much water flows past point Y, on average, in a month?

There are two important results in this case, also:-

A. The lake will begin to be filled up by sediments deposited in it; the coarser particles will form a delta at point D, and the finer particles will fall to the bottom in the almost still waters of the lake to form lacustrine sediments (LS).

B. At point Y the river will have a great deal of surplus energy because it will have a high gradient where it overflows the rim of the basin. It will erode its bed, lowering the level of the lake, causing it to shrink in area, and eventually (if the lake bed is above sea level) completely draining it.

This is how a river system is formed in an area of humid climate: one basin drains into another until eventually the surplus water finds its way out to the sea, and the basins are eventually eliminated.

N.B. In actual fieldwork E is extremely difficult to measure. It is possible to measure how much evaporation *would* take place from a water surface, or how much e/t from a given soil/vegetation combination, but to measure how much *has* taken place is well-nigh impossible in view of the variations of soil/vegetation, and in the availability of water for evaporation, amount of sunshine, wind speed etc. But Q is relatively easy to measure (see fieldwork example). The formula makes it possible to *calculate* e/t: it is not necessary to measure all the variables.

Answers

Case 1

Q.1 Total volume (P) = 37,500,000 m^3

Q.2 Total evaporation (E) on land = 30,000,000 m^3

Q.3 Surplus = 7,500,000 m^3

Q.4 Size of lake = 7.5 km^3

Q.5 Q = 0

Case 2

Q.6 P = 75,000,000 m^3
E/T on land surface = 15,000,000 m^3
extra E. on water surface = 3,600,000 m^3
Total E = 18,600,000 m^3
Q = 56,400,000 m^3 per annum.

Q.7 Q = 4,700,000 m^3 per month.

HYDROLOGY — FIELD WORK

This section is a reproduction of actual work undertaken in the valley of Great Langdale Beck Cumbria, in May 1979, including the results obtained.

STREAM FLOW

We wish to investigate how streams flow in their beds; we are interested because this controls the energy which they have available for the work of erosion, transportation and deposition. In this piece of field work we wish to investigate *two* things:-

> What is the relationship between the width and depth of a stream and its speed?
>
> How is the way in which a stream flows in a deep wide channel different from the way it flows in a shallow narrow one?

We shall also be able to make further observations about gradient and speed, and gradient and load.

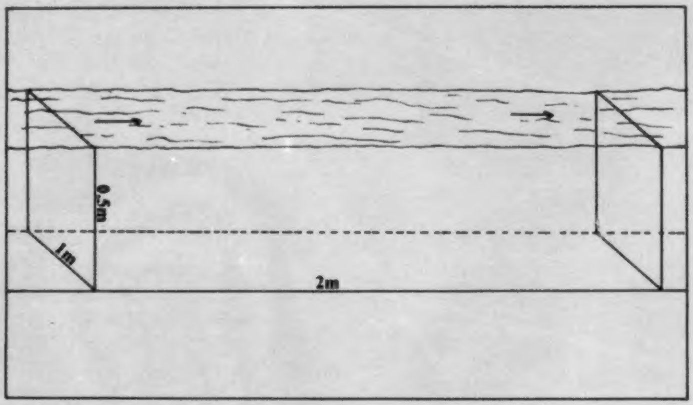


Fig. 16.

Measuring the flow of a stream

Look at Fig. 16. How would I find out the *volume* of water outlined in solid lines on the diagram?

Answer: multiply the width by the depth by the length.
2 metres × 0.5m × 1m = 1 cubic metre.

Supposing the stream was flowing fast enough for 1 metre to go past in 1 second: the rate of flow of the stream would be 1 cubic metre per second (shortened to 1 cusec).

Using this same stream as an example, what would be the flow of the stream if its speed were:-

a) 0.5 metres per second
b) 2.0 metres per second?

So we can see that the rate of flow of any stream is:-

Area of the cross section × speed of the current.

These are the things we need to measure if we are to find out how much water flows down a stream in a given time (per second, per day, per year).

(Think of how this could be useful to the following people: the physical geographer interested in the formation of landscape: the water engineer: the agricultural specialist in a Third World country: the Transport Minister of a continental country).

The following assignments are to be carried out at four points along Great Langdale Beck, by different groups of pupils. Careful records must be kept in the field, speed of flow and volume must be calculated in the laboratory in the evening, and the results presented in downstream order of stations used.

Assignment A Measure the area of the cross-section of the stream. This can be done by taking depth soundings at 1 metre intervals across the stream (see Fig. 17).

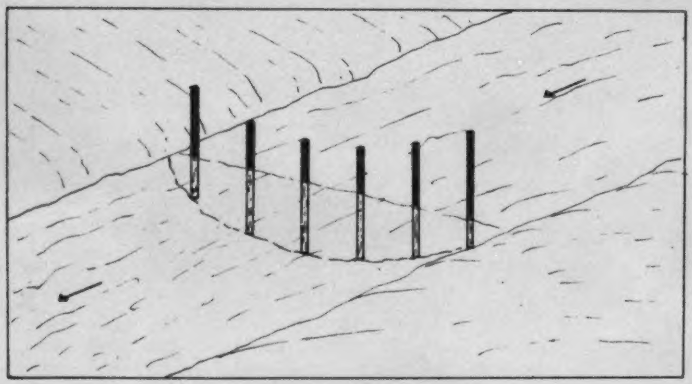


Fig. 17

Write down the readings in a table, like this:-

Distance from bank	Depth
1 metre	0.75m
2 metres	0.90m
etc.	

Then we can construct a cross section of the stream on a piece of graph paper, and measure the area.

Assignment B Measure the speed of the stream current.
Set up two ranging poles 10 metres apart. Have two pupils sighting across the stream at right angles to it, the upstream one with a whistle, the downstream one with a stopwatch. Throw an orange into the stream above the first pupil. When it passes the first post he blows the whistle, and the downstream one starts the watch, and stops it when the orange passes. That is the time it took for the orange to travel 10 metres: since it is approximately the same density as the water, and only just floats, the time is the same as the water takes to travel the same distance.

The speed of the stream varies across the surface, so we are more likely to get the average speed of the stream if we repeat the process several times at varying distances from the bank. So repeat the process 10 times, and take the mean of the 10 readings as t, the time taken for the stream to travel 10 metres at this point.

Assignment C Calculate Q, the flow of the stream at each of the four stations.

The distance travelled is 10 metres.
The time taken is t seconds.

The speed of the stream at the surface is $\frac{10}{t}$ metres per second.

For each station, multiply the speed of the stream by the cross-sectional area, and the result is the stream flow (Q.)

Assignment D Make notes on the behaviour of the stream at each site under these headings:-

(i) size and quantity of deposited load
(ii) type of flow (smooth or turbulent)
(iii) signs of downwards or sideways erosion in the stream bed.

Hydrology Field Work - Results

Assignments A B and C were carried out by a party of 14-year old girls on a field expedition to the Langdale Valley, in Cumbria. No difficulty was experienced with the taking of readings, but help was needed on the calculations, and on their meaning.

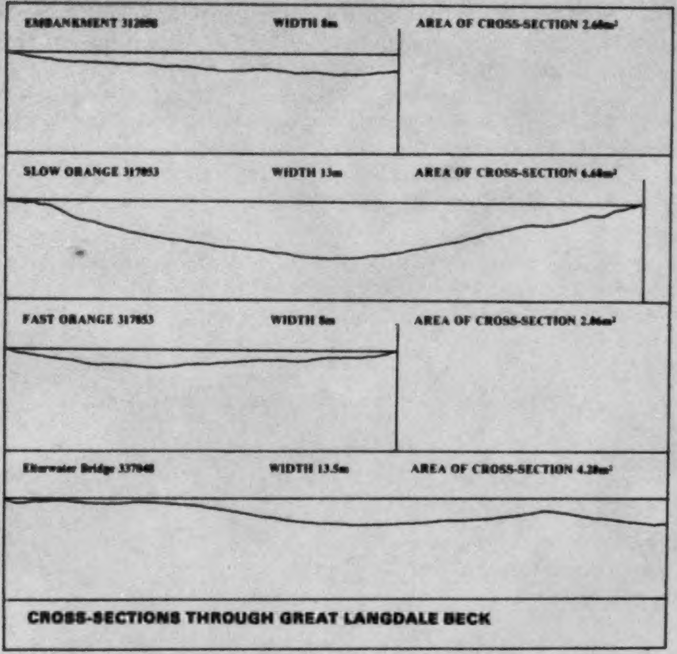

EMBANKMENT 312098
WIDTH 8m
AREA OF CROSS-SECTION 2.60m²
SLOW ORANGE 317053
WIDTH 13m
AREA OF CROSS-SECTION 6.60m²
FAST ORANGE 317053
WIDTH 8m
AREA OF CROSS-SECTION 2.06m²
Elterwater Bridge 337048
WIDTH 13.5m
AREA OF CROSS-SECTION 4.20m²
CROSS-SECTIONS THROUGH GREAT LANGDALE BECK

Fig. 18.

Fig. 18 shows the cross-section drawn from the readings and the cross-sectional area taken off the diagram.

The table below shows the numerical results obtained:-

	Area of stream cross-section in m²	Time taken for the stream to flow 10 metres	Speed in metres per second	'Q'-Stream flow in cubic metres per second
1	2.60	27.0	0.37	1.04
2	6.60	57.4	0.17	1.122
3	2.06	12.2	0.82	1.69
4	4.20	15.0	0.61	2.56

1. 'Embankment'
2. 'Slow orange'
3. 'Fast orange'
4. Elter-water Bridge.

A. The results were reasonable, since the figures for Q come out in the right order – increasing volume downstream.

B. The result for Q at stations 2 and 3 show the limitations of the simple techniques used: since they were next to each other, the volume should theoretically be the same.

C. However, stations 2 and 3 illustrate clearly the two ways in which a stream can flow at approximately the same volume:
   - slowly as a wide stream, as in 'slow orange'
   - fast as a narrower, shallower stream as in 'fast orange'.

D. The work could be usefully extended by accurate measurements of stream gradient, but this is difficult with the resources usually available.

AN OUTLINE OF SYLLABUS AREAS THAT HAVE LINKS WITH MATHEMATICS

GEOGRAPHY	— LINKS WITH MATHS
HUMAN GEOGRAPHY Key idea is the relationship: area – line – point. Concentration and Dispersal. All man's activities can be seen in terms of this relationship (See Ex 1) and Settlement below.	
Agriculture Two views:- A. Areal – progress from peasant subsistence farming to a world-integrated commercial farming economy seen as an increasing ability to overcome the resistance in the system. So *transport* seen at an early stage in terms of the 'friction of distance' by reconstructing the modes of transport applicable to primitive economies. Then the idea of 'lubrication' reducing this friction, by the development of transport can be introduced.	Development of the market-area polygon of Christaller first in the case of the subsistence farmer (Example 1) – then the concept is available for use in studying settlements
– development of farming zones of varying type and intensity (von Thunen) around a town, and varying land-value. (Both these seen in conjunction with the growth of settlement, and using the same concepts).	Line graphs converting into a circular/zonal pattern – in concepts of 'economic rent' and bid-rent curves
B. Systematic – Agriculture seen as a system with inputs, outputs, throughputs, feed back, etc. with a 'random' element.	Flow-line diagram with feedback loops. Simple quantitative data fed into the model to illustrate alternative inputs (eg Capital/machinery to replace labour) and outputs (subsistence versus sale of industrial crops) etc.
The interaction of both indicated; change in intensity of farming changes the *scale* of the area model, changes in output change the areal patterns (of food procurement, raw material procurement etc)	Calculations putting transport/distance data into numerical models link the two approaches

GEOGRAPHY	*LINKS WITH MATHS*
Industry Again Two Views:-	
A. Areal – The major difference from the agriculture model is the 'sporadic' distribution of many industrial raw materials such as mineral deposits.	Weber's Material Index introduced at an early stage. Weber's mechanical model useful *in a historical context*.
NB. Progression (paralleling changes in agriculture) from universally available raw materials in primitive craft industries, to increasingly sporadic materials:- wood, clay – copper, iron, petroleum, titanium.	Market area concept linked to area/point theory and Christaller as in Ex 1.
The key again is the improvement in transport technology lubricating the 'friction of distance'. Liberation of industry from the 'tyranny of raw materials' by improved transport switches the control from the realm of Physical Geography to that of Human Choice.	
B. Systematic – the same input/output/feedback model can be used as for agriculture (with modifications). The random element changes from the 'Act of God' to variations in human taste and choice. The role of capital accumulation is seen as more important, in view of its greater freedom to generate wealth in the industrial field, freed of many of the inherent restrictions seen in agriculture (land tenure systems, soil fertility, climatic limits).	Theoretical figures fed into the model to illustrate the effect of:- alternative inputs of – capital and labour (mechanisation) – increased production leading to lower unit cost (economies of scale) – different rates of reinvestment (problems of use of resources in Third World countries).
The systematic model is considered first in relation to one firm, to illustrate the role of profit, capital accumulation and investment and growth. Then the concepts are applied to the problems of industrial investment in under-developed countries.	
Interaction of the two views leads to an appreciation of the development of industry from dispersed small-scale craft industries to modern large-scale integrated factory industry.	The concepts needed arise from the industrial and agricultural models above, and the spatial/economic concepts of town development as below.
This development seen as parallel to the development of towns.	

GEOGRAPHY	LINKS WITH MATHS
Centralisation of towns and industry posed as a problem, and the advantage (high material standard of living) and the disadvantages (vulnerability to political/social unrest, high rate of exhaustion of world resources, high rate of pollution) discussed. The *mechanisms* of industrialisation and urbanisation should be clearly understood after the presentation of these concepts, allowing for discussion of the politics, economic, social and moral choices in an informed way.	Statistics of population and economic patterns introduced at an early stage – pupils must be familiarised with typical rural/urban population figures and the increasing concentration in towns, *in statistical form.*
Settlement	
Function the key to understanding. Central Places providing general services of different types and levels discussed in terms of Christaller's model. Range and Threshold introduced and clarified at an early stage. The rank-size rule, and the correlation of settlement size with the number of retail outlets (or any other service) investigated. Specialist settlements of various types discussed, and the statistical methods of distinguishing them: industrial settlements (employment statistics), transport oriented settlements (traffic statistics), retirement settlements (employment figures, Health Authority figures) etc.	The Christaller landscape (which, if the sequence is as here, will be familiar by now) is the basis for the spatial concept. The effect of increased personal mobility, and improved methods of storing food on the *scale* of the Christaller model indicated – and the question of the continued validity of the model discussed. Simple ideas of correlation of variables (by graphs) and the investigation of 'abnormal' cases revealed.
Internal structure of towns – Functional zones, density variations, increasing differences with increasing mobility and social differentation in area.	Bid-rent curves, turned into zonal variations. Ditto density-distance graphs. Land value 3D 'surfaces' shown (no attempt at construction) Density/area calculations. Use of census material to define sub-standard housing.
Further analysis of the effect of the motor car on town structure and function, expansion of town area, 'non-functional' land uses, etc.	
Growth of 'Megalopolis' enabled and encouraged by evolutions already discussed in the agricultural/industrial sphere. General overview of integration and concentration forces, the problems caused, and the beginnings of dispersal movements in individual behaviour and Government policies.	
The structure of society seen in terms of economic forces, and the limits of political and personal choice.	

GEOGRAPHY	LINKS WITH MATHS
Transport	
Transport is the key to understanding most of the ideas in economic development – the idea of the 'friction of distance' and its reduction by technical developments. So its discussion cannot be reserved for a separate section, but must be taught along with the changes being discussed in agriculture and industry.	Network theory, shortest-path matrices, centrality etc. New links fed into a mathematical model of specific areas to predict their effect on the settlement/transport pattern. See Ex. 2.
This leaves areal patterns to be discussed in detail, and to a large extent these too are imposed by the settlement and economic factors – what is to be transported, and where. Transport patterns, and ideas of centrality and accessibility crucial to the ideas of town growth – ideal locations and (much more frequently) eccentric locations. The use of an ideal pattern is here (as frequently elsewhere) much more useful in focussing ideas about the departure from the ideal, than in explaining what does exist. Thus the excentric location of main cities in colonial countries, and the attempts by governments in Brazil and Australia (for example) to counteract these tendencies by building new inland capitals are more clearly seen.	Concept of the gravity model (not used in its mathematical expression before the Sixth Form) and the 'breaking-point'. The effect of transport on the gravity model and its role in by-passing smaller settlements. 'Zarb' example.
The positive feedback model of: Transport Advantages – leads to – siting and growth of a main town – leads to the building of more transport – leads to further growth etc. Paris taken as the best example, leading to the question (as above) of the limits of centralisation, the attempts at dispersal, and the fostering of greater regional autonomy.	
Population	
Areal View – density and distribution related to economic level to attempt to define 'optimum population' and 'overpopulation'. Distribution seen as a function of the above sections on agriculture and industry, and the increasing economic integration of a single country, and the World.	Calculations of mean densities, and their limitations (eg in Egypt). Age/Sex pyramids seen as a special form of bar graph (cf. wind rose).

GEOGRAPHY	LINKS WITH MATHS
Systematic View – Birth Rate, Death Rate, Natural increase, age structure. Simple view of the population development model and the population explosion. Effects of various rates of increase, projections according to different assumptions, limits to growth (Malthusian theories) etc. Population and economic levels.	Correlation of economic information (GNP per capita) with demographic – birth, death and natural increase rates, expectation of life, etc. Use of set theory to attempt explanations.

PHYSICAL GEOGRAPHY

Hydrology

The Water Cycle seen in terms of WATER The question 'What happens to rain' taken as the starting point for an analysis of stream flow, runoff rates, underground storage of water, evaporation rates etc., seen first in terms of *water*.	Precipitation:- (P) Calculation of *volume* given area of a river basin and height of precipitation. Calculation of river flow (Q) from speed of flow and cross-sectional area. Difficulty of calculating actual rates of evaporation (E) Use of the simplified model $Q = P - E \pm \Delta S$ Demonstration how the formula eliminates need to measure all variables. (Examples 3 and 4) Use of storm hydrograph and *average* figures to illustrate function of Storage (S) as a regulator; serious consequences of interfering with S.
The system of salt concentration in seas, the rate of evaporation in enclosed basins – concepts of inland drainage basins, and the evolution of integrated drainage systems.	Different rates of evaporation seen as controlling Q Evaporation from open water surfaces increasing with area. Zero values of Q.
Glaciers seen as having the same function as rivers. Comparison of cross-sectional areas and rates of flow. % of surface water tied up in each case. Effect on sea levels.	Volume and flow calculations.

GEOGRAPHY	— LINKS WITH MATHS
The Water Cycle seen in terms of *ENERGY* Input of solar energy (refer to Meteorological aspects). Output in the form of kinetic energy – total energy available depends on volume of precipitation and height of land on which it is deposited.	Actual energy calculations seen as inappropriate to school Geography. Gradient calculations linked to ideas of tangent.
Rate and manner of loss of the energy seen in terms of gradient, efficiency of stream bed and size (turbulent and smooth flow)	Rate of loss at different gradients seen in terms of % loss per km.
'Excess energy' used in transporting, eroding and reducing the particle size of its load. Effect of changes in energy – erosion, deposition, changes in the river's capacity and competence etc. (In other words the Dynamic Equilibrium concept.)	Again (as throughout) correlation of two variables by simple graph construction used. Here actual figures obtained in field work used where possible – eg variation of particle size measured at various points along a river plotted against speed of current.
Meteorology/Climatology	
Temperature and Precipitation figures (Preferably based on direct observation in school) and the instruments used.	Mean median mode Simple ideas of deviation Use of line and bar graphs, wind roses, etc.
Relative and absolute humidity, Dew Point, variation of RH with cooling.	Use of % change and distinction from absolute change.
Energy budget of the atmosphere studied in % terms, differential heating of land/sea/snow surfaces, effect on diurnal/annual temperature ranges.	Use of simple formulae, to stress heat as something which does not disappear, and can be accounted for.
Vertical stability of the atmosphere explained in simple terms.	
Variations in the 'budget' at different latitudes leading to import and export of heat – Hadley Cell, ocean currents.	Slightly more complex formulae bringing in the possibility of regional exchange of heat, storage in the form of plant tissue etc.
Climatic effect of man on the Meteorological aspects of the water cycle (forest clearance leading (?) to lowered rainfall, cloud seeding,) and on the climate in towns and cities (higher night-time temperatures, funnelling of winds etc.)	

ICSU Committee on the Teaching of Science

Edited by Alan Rogerson

③

Our Inheritance: Common Ground for the Mathematics and Biology Teacher

Brian Dudley

∫ydx
$\frac{-b \pm \sqrt{b^2-4ac}}{2a}$
010
001
×
÷
%
C_2H_5OH
$E=mc^2$
NaOH

Acknowledgements

We are most grateful to the Institute of Biology for permission to use materials from a number of papers first published in the Journal of Biological Education.

One of the most striking transformations during the past twenty years has been the rapid development of the relation between mathematics and biology. A sudden realisation has occurred that a vast field of previously unexploited data is available for mathematical modelling. In the past biology has been largely a descriptive science with a conspicuous lack of interaction with mathematics. There is now an extensive interaction of biology with various branches of mathematics: graphs, algebra, calculus, differential equations, vector analysis and statistics.

This booklet is unusual in this series in that it develops *one* theme to show the use and development of mathematical ideas and terminology in a science – in this case biology. Much of the 'new' biology is now integrated with mathematics, but this is not always clear or accessible at school level. This booklet develops an important biological theme – genetic inheritance – and in the process shows the use and relevance of mathematical ideas such as probability, cartesian product, tree diagram, set notation etc., as well as the more familiar numerical work. All of this material is accessible up to and including VIth form students (16-18 year olds) in Britain, and this case study should be of great interest to both mathematics and biology *teachers*. This is just the tip of the iceberg, however, and local cooperation between maths/biology teachers will soon reveal enormous areas of common interest and support, in their respective subjects. It's up to you!

Our Inheritance: Common Ground for the Mathematics and Biology Teacher

One important exercise in biology is the classification of information (that is, the identification of set and subsets). This can be seen most readily considering living things. For example insects, crabs, spiders, centipedes and the like all belong to one *phylum*, the Arthropoda (literally, animals with jointed legs); within this is the *class* of insects, the Insecta, of which one of the many *orders* is the Hymenoptera (wasps, ants and bees). In this classification the biologist is concerned to partition the set and subsets, while at the same time raising a succession, or hierarchy, of sets and subsets which reflect what is believed to be the natural evolution of the group in question, the so-called 'natural classifications' of biology. Both concerns can give rise to lively debate! (See Dudley 1977a, Chapter 3).

Other examples of sets are often found in biology. One includes the set of genes of an individual, the set of parents and the set of all possible types of fertilized egg (zygote) which can arise from them. In taxonomy, which is the science of classifying animals and plants, the "elements" of a set are whole organisms (a man, a fern, a grass) but in genetics, while some sets also may be made up of individual organisms, others are composed of individual genes, individual zygotes and so forth.

The identification of sets, sub-sets and their elements and the relations between them is also of course part of mathematics. Sets can be represented by letters and relations and operations by symbols. This method of notation has been used to portray classification in biology (Reid and Booth, 1970, and Dudley, 1977 (a), (b)) and blood groupings (Rebman, 1967) and also can be used to represent sets in genetics. For example, the set of parents, P, of a single individual is made up of a male parent, c, and a female parent, d, as follows:

$$P = \{c, d\}$$

Any gene, *h*, of a set of genes (H) of an individual is itself a set made up of two elements – *m*, contributed by the male parent and *f*, contributed by the female parent. It is these elements, *m* and *f*, which the biologist calls "alleles".

An individual is made of many cells all of which arise from the one fertilized egg, or zygote. With the important exception of one special type of cell – the gamete – all cells of the individual have two elements (alleles) to each gamete (see Figure 1). Each allele is part of a chromosome and the two alleles that make up a gene are on two similar chromosomes which make up a pair of 'homologous' chromosomes. The alleles for different genes may be on different homologous chromosomes, or they may be on the same homologous pair when they are said to be linked.

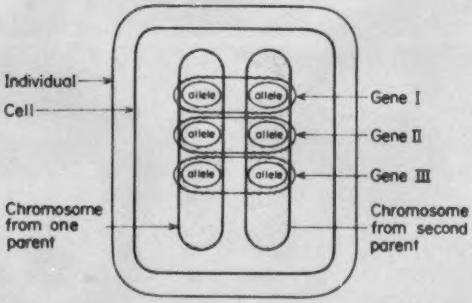


Figure 1. *Diagrammatic representation of the genetic composition of an individual.*

When a male gamete (a sperm in the case of animals, the inside of pollen grains in the case of flowering plants) fuses with a female gamete (an ovum, or unfertilized egg) the fertilized egg, or zygote, results in which each gene is again made up of two alleles. The individual grows from the zygote. (Figure 2)

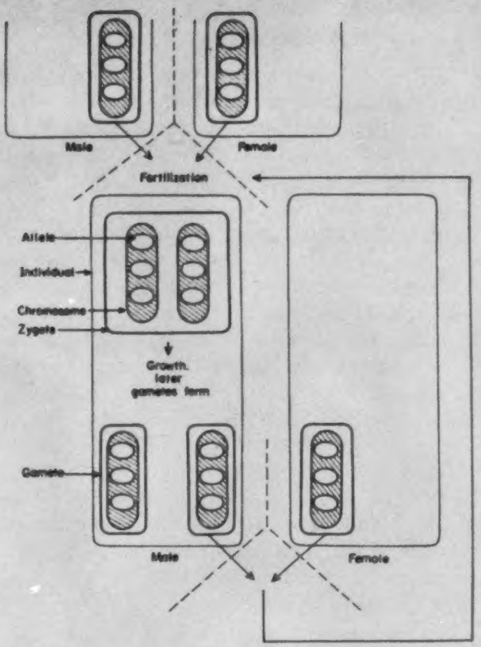


Figure 2. *The cycle of the inheritance of genetic material.*

It is precisely because the maternal and paternal chromosomes do not fuse but remain discrete within the fertilized egg that each gene of an individual is made up of a pair of elements (alleles), one in each of two chromosomes. As a consequence a gene, h, of the set of genes, H, consists of a pair of alleles:

$$h = \{ m, f \}$$

where m is an allele of gene h that has been inherited from the male parent, and f is an allele of gene h that has been inherited from the female parent. Every gene is made up of selected pairs of elements, one from each of two sets, namely the set of alleles (M) of the male parent and the set of alleles (F) of the female parent.

To sum up:

1. A fertilized egg, and hence the individual arising from it, originates from one male gamete fusing with one female gamete.
2. Each gamete contains only a single allele (element) of each gene.
3. Each individual (organism) contains a pair of alleles (elements) for each gene.
4. The number of chromosomes per cell varies between species, for example it is 46 for humans, 308 for mulberry trees (White, 1942), while for one particular fern it is approximately 1020 (Manton and Sledge, 1955).

The inheritance of a single character

Biologists work from the premise that each character or feature of the body is controlled by a single gene of two alleles. While readily acknowledging that this premise is only part of the known truth concerning inheritance, it is quite sufficient for our present purposes, and will be adopted here. The study of the inheritance of such characters in isolation is called "monohybrid inheritance" though there is more than one form to the character taken for study. For instance, the one gene for eye colour in man may result in the eyes being either brown or blue, the alternative colours resulting from the different alleles of the one gene. One of these alleles (all the evidence leads us to suppose it to be for brown eyes in this instance) is expressed whenever it is present, while the other (for blue eyes) is expressed only when the first type is absent. The first type therefore is said to be *dominant* and is represented by a "capital" - or upper case - letter while the other allele is said to be recessive and is represented by the corresponding "small" - or lower case - letter.

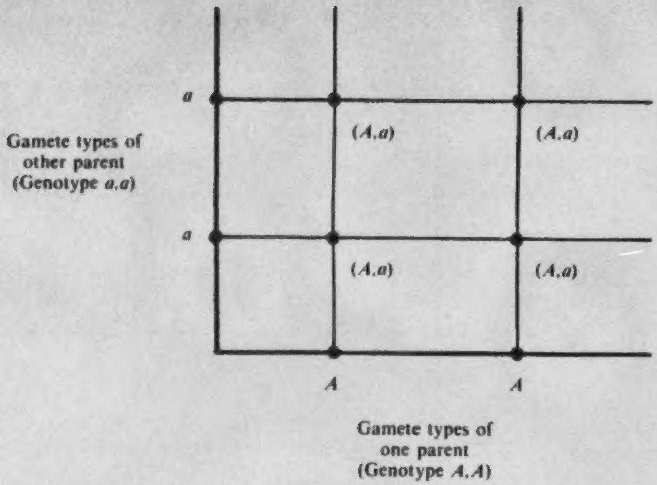
Even this rule does not fit all known instances concerning the inheritance of eye colour in humans. Exceptions exist, as the data from one Danish survey shows (Table 1), and the exceptions - the grey-green eyed children and the blue-green eyed children are difficult to interpret. Again, in spite of these known complexities, the rule applies perfectly in so many instances that it will be taken as true for our present purposes.

Table 1: Data on eye colour of parents and their children (from a Danish survey, quoted in 'The human pedigree' by A. Smith, 1975, Allen and Unwin, page 107).

Parents eye colour		Expected eye colour of offspring	Number of children		
Mother	Father		Blue eyes	Brown eyes	Grey/Green or Blue/Green eyes
Blue	Blue	All Blue	625 (97.0%)	12 (1.8%)	7 (1.1%)
Brown	Blue	Blue or Brown	317 (48.9%)	322 (49.7%)	9 (1.4%)
Brown	Brown	Blue or Brown	25 (23.4%)	82 (76.6%)	0 (0.0%)

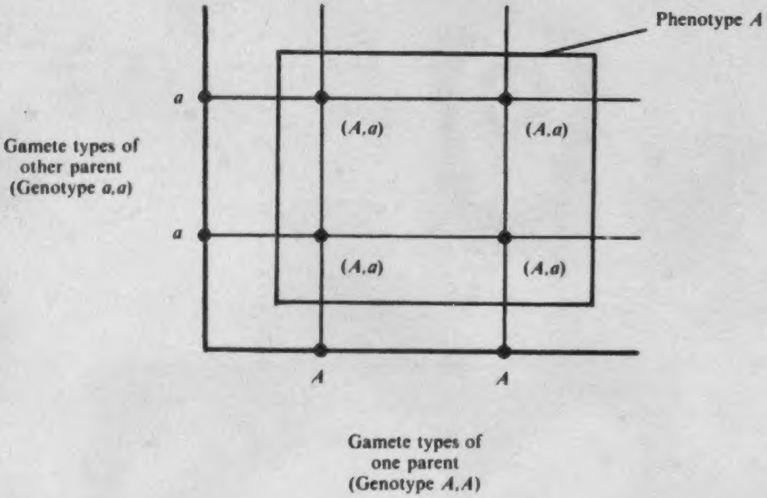
Table 1

Consider an arbitrary gene *A*. When an individual has both alleles of this gene dominant then its genetic composition (its *genotype*) is represented by *A,A*. When an individual has both alleles of this gene recessive its genotype is represented by *a,a*. When one parent has the genotype *A,A* and the other *a,a*, the types of offspring possible can be shown by the co-ordinates of a cartesian lattice, where the types of possible genotypes amongst the offspring are shown by the intersections thus:

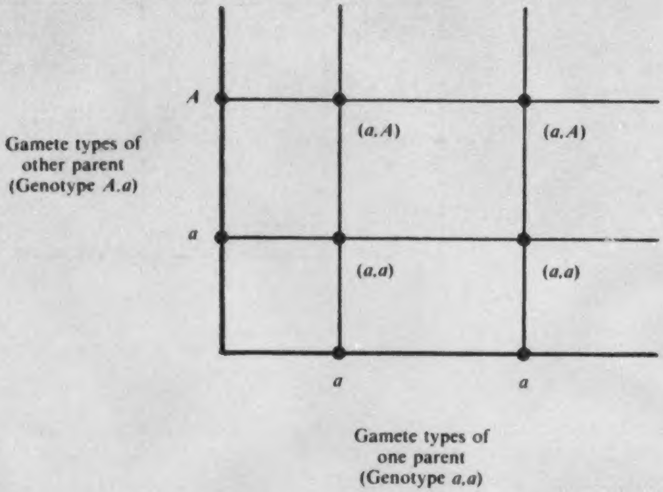

a
(A,a)
(A,a)
Gamete types of
other parent
(Genotype a,a)
a
(A,a)
(A,a)
A
A
Gamete types of
one parent
(Genotype A,A)

This arises because the alleles of a single gene appear in separate gametes; only one allele of a gene occurs in any one particular gamete, as mentioned earlier.

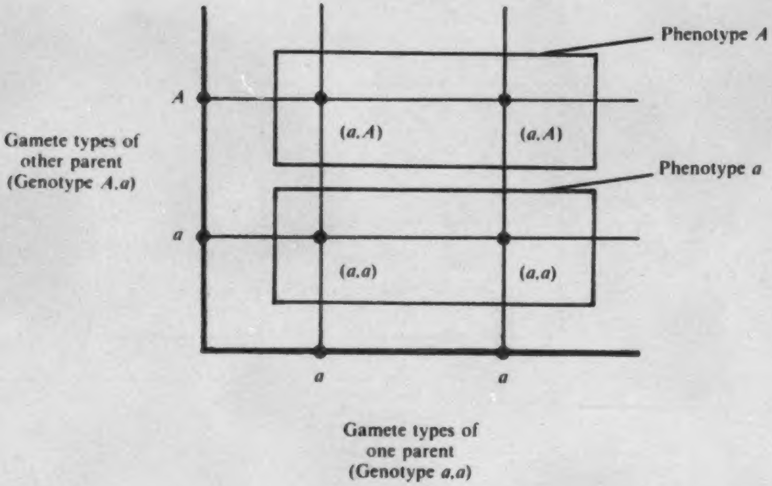
All the offspring in this example are of the genotype (*A,a*) and have the appearance (that is, the *phenotype*) corresponding to dominant allele *A*. Indeed all the offspring will, so far as this gene is concerned, have the same appearance as did the first parent. The phenotypes can usefully be enclosed. The instance above, then becomes


Phenotype A
a
(A,a)
(A,a)
Gamete types of
other parent
(Genotype a,a)
a
(A,a)
(A,a)
A
A
Gamete types of
one parent
(Genotype A,A)

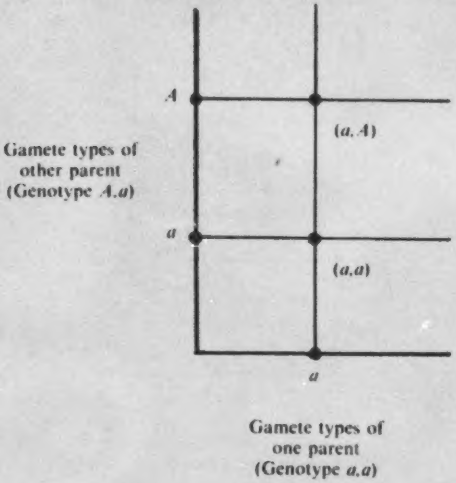
When one parent has the genotype *a,a* and the other *A,a* the types of offspring that are possible are shown to be as follows


A
(a,A)
(a,A)
Gamete types of other parent (Genotype A,a)
a
(a,a)
(a,a)
a
a
Gamete types of one parent (Genotype a,a)

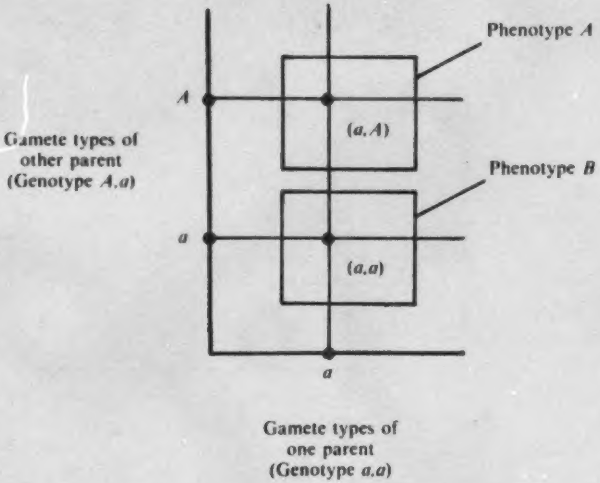
When the genotypes that give rise to the same phenotypes are enclosed this time, the result is


Phenotype A
A
(a,A)
(a,A)
Gamete types of other parent (Genotype A,a)
Phenotype a
a
(a,a)
(a,a)
a
a
Gamete types of one parent (Genotype a,a)

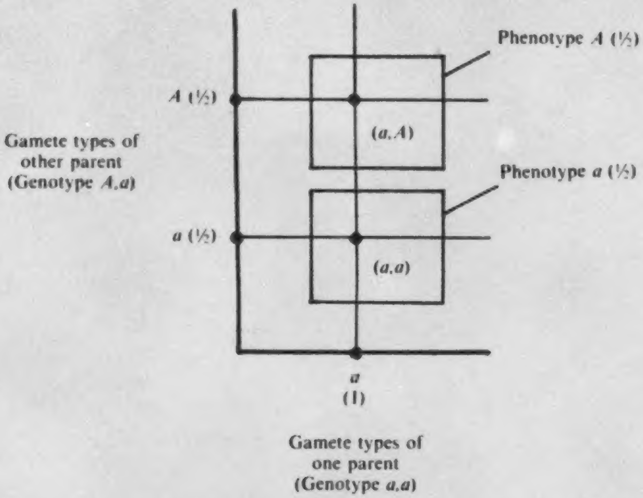
Here two kinds of product can form, namely (A,a) and (a,a). On this occasion the subsets of the set of gametes produced by the recessive parent, (a,a), are identical and so they may be united though those of the other are different and so may not be united. The simplified cartesian lattice then appears as


A
(a, A)
Gamete types of
other parent
(Genotype A,a)
a
(a,a)
a
Gamete types of
one parent
(Genotype a,a)

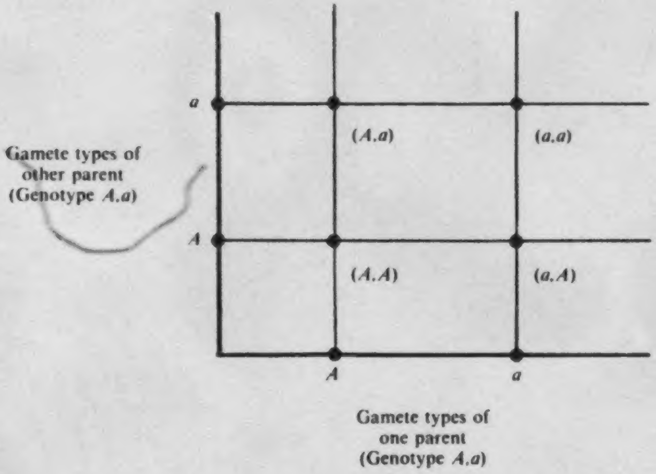
and, with the enclosure of genotypes giving rise to identical phenotypes,


Phenotype A
A
(a, A)
Gamete types of
other parent
(Genotype A,a)
Phenotype B
a
(a,a)
a
Gamete types of
one parent
(Genotype a,a)

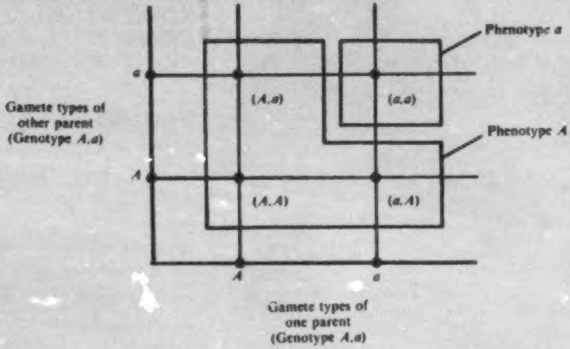
Experiments that have been performed with parents of this type (namely the cross described in biology by $A,a \times a,a$) show not only that two types do occur amongst the offspring, but also that they occur with approximately equal frequency, provided the size of the population of offspring is large. In consequence, the relative frequency of the gamete types *must* be as follows:


Phenotype A (½)
A (½)
Gamete types of other parent (Genotype A,a)
(a,A)
Phenotype a (½)
a (½)
(a,a)
a
(1)
Gamete types of one parent (Genotype a,a)

It is also possible for both parents to have the genotype A,a, when the possible genotypes of the offspring are as follows:

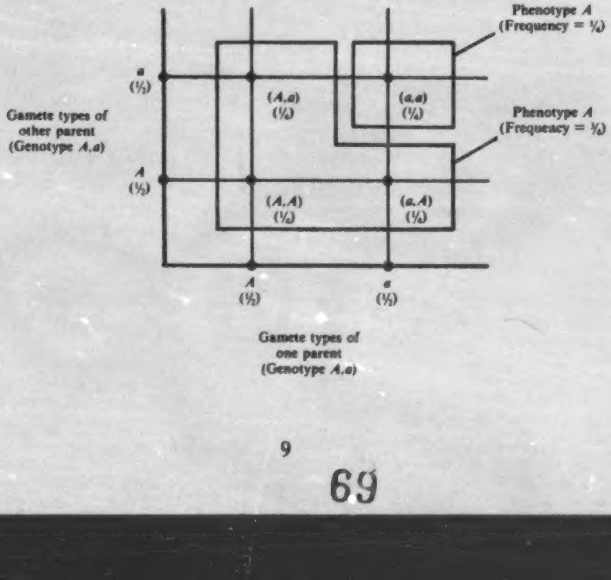

a
(A,a)
(a,a)
Gamete types of other parent (Genotype A,a)
A
(A,A)
(a,A)
A
a
Gamete types of one parent (Genotype A,a)

and enclosures of the phenotypes of the same kind results in the following:


Phenotype *a*
Phenotype *A*
a
A
(*A*,*a*)
(*a*,*a*)
(*A*,*A*)
(*a*,*A*)
Gamete types of other parent (Genotype *A*,*a*)
A
a
Gamete types of one parent (Genotype *A*,*a*)

Four kinds of genotype result amongst these offspring, though in genetics (*A*,*a*) is always functionally indistinguishable from that of (*a*,*A*). In three of the four possible types the dominant allele, *A*, is present. In only one are there recessive alleles alone, namely, (*a*,*a*). Thus amongst the offspring arising from this cross there will be individuals of two kinds (phenotypes), some having the phenotype of the dominant allele (phenotype *A*) and some that of the recessive allele (phenotype *a*).

Mendel, with his work on peas, showed that the dominant and the recessive phenotype individuals occur in the ratio of 3 to 1 provided the size of the population of offspring is large, and this ratio can be explained only by assuming once again that the gamete types of each parent are equally frequent; that in each parent the frequency of gamete type *a* is equal to that for gamete type *A*. (= ½). The frequency of the sub-set of offspring identified as having the recessive phenotype *a* must be ¼, giving the classical 3 : 1 ratio as it is written in biology. In diagram form these frequencies and the ratios that result from them are shown as follows:


Phenotype *A* (Frequency = ¼)
Phenotype *A* (Frequency = ¾)
a (½)
A (½)
(*A*,*a*) (¼)
(*a*,*a*) (¼)
(*A*,*A*) (¼)
(*a*,*A*) (¼)
Gamete types of other parent (Genotype *A*,*a*)
A (½)
a (½)
Gamete types of one parent (Genotype *A*,*a*)

The experiment in which both parents are A,a (namely the genetic cross designated $A,a \times A,a$) and the possible offspring that can be formed from it can also be written as follows

Set of allele types of the one parent is $\mathrm{M} = \{A,a\}$

Set of allele types of the other parent if $\mathrm{F} = \{A,a\}$

Set of possible types of offspring is

$$\mathrm{M} \times \mathrm{F} = \{(A,A), (A,a), (a,A), (a,a)\}$$

Similarly in the experiment represented earlier as $A,a \times a,a$,

Set of allele types of the one parent is $\mathrm{M} = \{A,a\}$

Set of allele types of the other parent is $\mathrm{F} = \{a,a\}$

Possible types of offspring are

$$\mathrm{M} \times \mathrm{F} = \{(A,a), (A,a), (a,a), (a,a)\}$$

Which is the set $\{(A,a), (a,a)\}$

In both, and all such cases of monohybrid inheritance, M × F is the "cartesian product" of the sets M, F. Mendelian inheritance operates as the cartesian product.

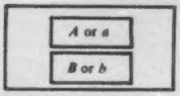
The inheritance of two characters

The study of the inheritance of two genes simultaneously is called dihybrid inheritance. In such studies it is first necessary to identify the types of gametes produced by each parent in the experiment and this can be a matter of confusion amongst biology students because the model which is commonly used to represent it is incomplete.

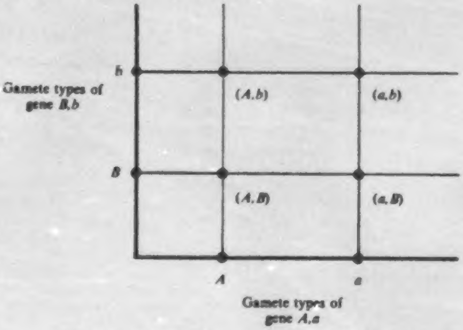
One classical experimental situation is where one parent possesses only the dominant alleles of the genes being considered while the other parent posseses only the recessive alleles of these genes. If the two genes under consideration are given the symbols A and B, then such an experiment is traditionally represented by the notation $A,A,B,B, \times a,a,b,b$. This possibly confusing notation simply means that the gamete from the one (dominant) parent has one allele from each gene and hence one allele A and one allele B, while the gamete from the other (recessive) parent, also necessarily with one allele from each gene, will have one of type a and one of type b. So (B,A) or (A,B) matches with either (a,b) or (b,a) and the possibilities for the genotypes of the offspring are (A,B,a,b), (B,A,a,b), $(B,A,b,a,)$ and (A,B,b,a); all the offspring must be genetically identical and all of the type A,a,B,b.

In appearance (phenotype) they will be of the type represented by A and B and will all appear similar to the dominant parent (whose phenotype was also A,B).

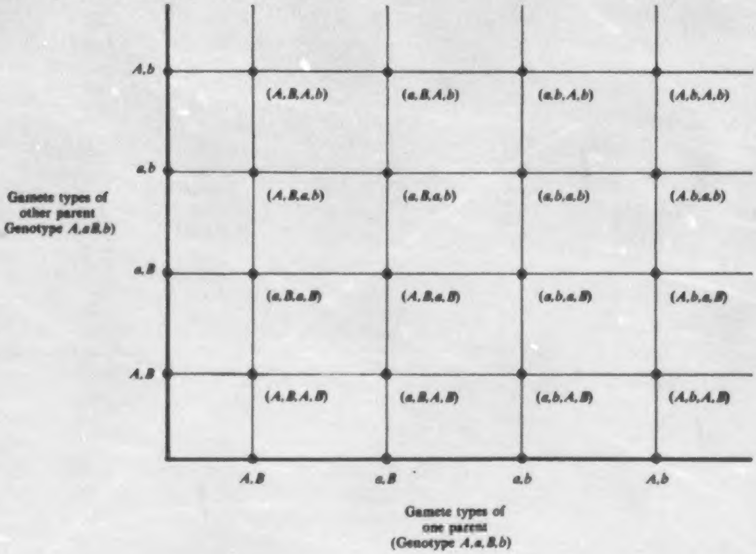
When these offspring, namely with genotype *A,a,B,b*, (but phenotype *A,B*), themselves function as parents and breed amongst themselves, then each produces gametes of the types



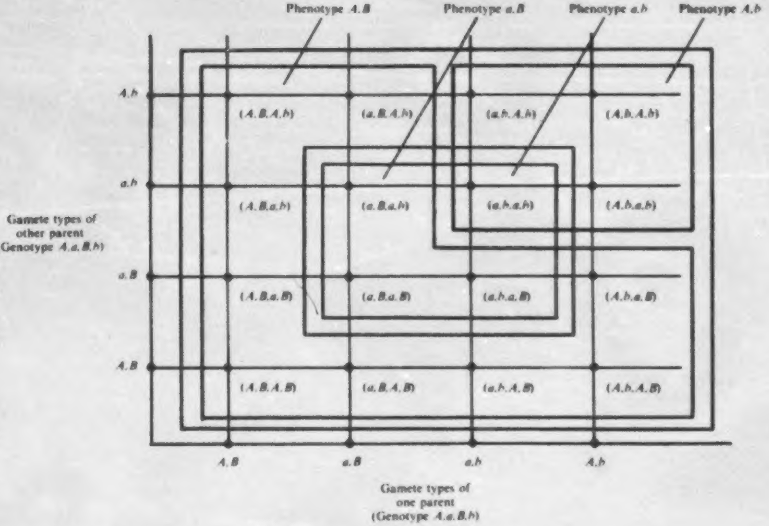
and the types of *gametes* produced by the parent *A,a,B,b*. can be found by means of the cartesian lattice as follows:-


b
B
Gamete types of gene B,b
(A,b)
(a,b)
(A,B)
(a,B)
A
a
Gamete types of gene A,a

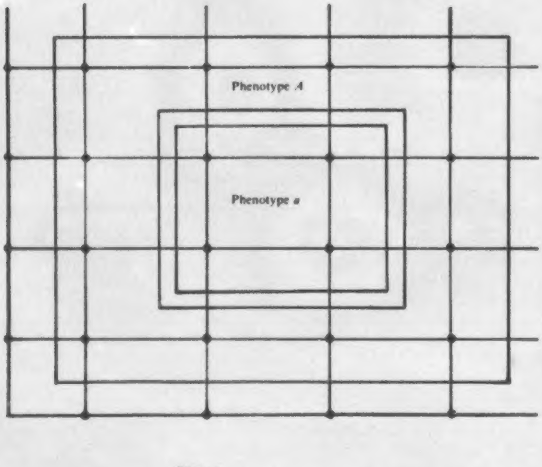
Each parent manufactures all four of the kinds possible. The kinds of offspring that result from the experiment *A,a,B,b* × *A,aB,b* can also be shown on a cartesian lattice:


A,b
a,b
a,B
A,B
Gamete types of other parent Genotype A,aB,b)
(A,B,A,b)
(a,B,A,b)
(a,b,A,b)
(A,b,A,b)
(A,B,a,b)
(a,B,a,b)
(a,b,a,b)
(A,b,a,b)
(a,B,a,B)
(A,B,a,B)
(a,b,a,B)
(A,b,a,B)
(A,B,A,B)
(a,B,A,B)
(a,b,A,B)
(A,b,A,B)
A,B
a,B
a,b
A,b
Gamete types of one parent (Genotype A,a,B,b)

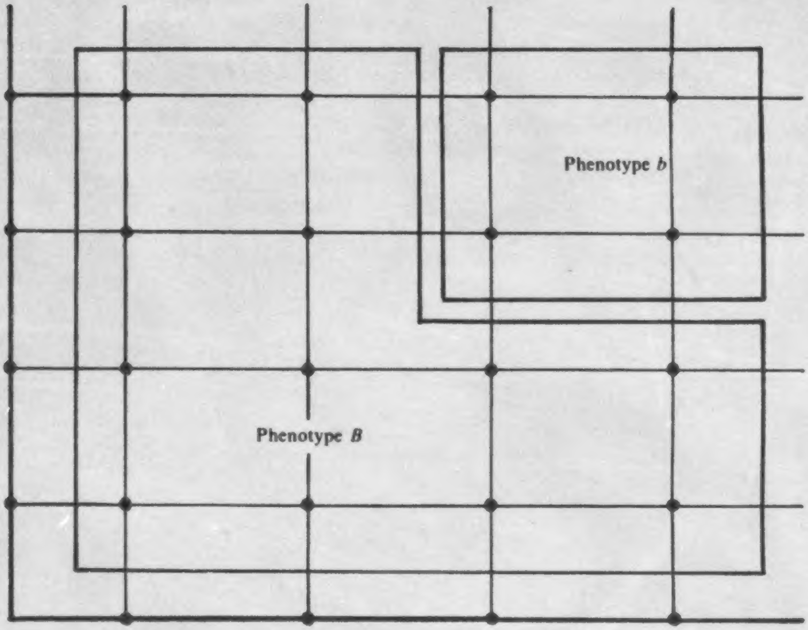
The phenotypes that can be formed amongst the offspring in this experiment are shown below.


Phenotype A.B
Phenotype a.B
Phenotype a.b
Phenotype A.b
A.b
a.b
a.B
A.B
(A.B.A.b)
(a.B.A.b)
(a.b.A.b)
(A.b.A.b)
(A.B.a.b)
(a.B.a.b)
(a.b.a.b)
(A.b.a.b)
(A.B.a.B)
(a.B.a.B)
(a.b.a.B)
(A.b.a.B)
(A.B.A.B)
(a.B.A.B)
(a.b.A.B)
(A.b.A.B)
Gamete types of other parent (Genotype A.a.B.b)
A.B
a.B
a.b
A.b
Gamete types of one parent (Genotype A.a.B.b)

The lattice shows that $\frac{3}{4}$ ($\frac{12}{16}$) of the offspring types have the phenotype *A* and $\frac{1}{4}$ ($\frac{4}{16}$) have the phenotype *a*. In simplified form, these show in the lattice as follows.


Phenotype A
Phenotype a

Similarly the lattice shows that $\frac{3}{4}$ ($\frac{12}{16}$) of the offspring types have the phenotype *B*, and $\frac{1}{4}$ ($\frac{4}{16}$) have the phenotype *b*. In simplified form, these show on the lattice as follows



Both genes, taken separately, give the same result as that for monohybrid inheritance and this is as one would expect.

When both genes are considered simultaneously, four phenotypes can be identified as subsets of the lattice, namely (*A*, *B*), (*A*, *b*), (*a*, *B*) and (*a*, *b*). These also derive from the cartesian product of the phenotypes that can arise from each gene.

Set of phenotypes of the one gene is $P = \{A, a\}$

Set of phenotypes of the other gene is $Q = \{B, b\}$

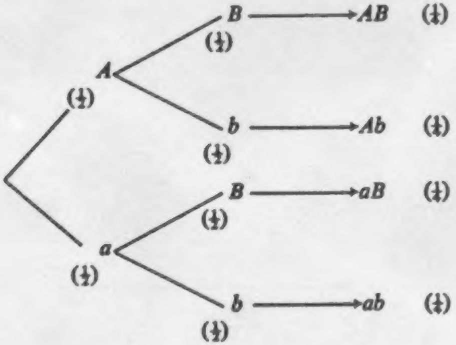
Set of possible phenotypes amongst the offspring is

$$P \times Q = \{(A,B), (A,b), (a,B), (a,b)\}$$

However these four do not occur with equal frequency amongst the types of possible offspring shown in the lattice for the experiment *A*,*a*,*B*,*b* × *A*,*a*,*B*,*b*.

There one finds that $^9/_{16}$ of all possible types have phenotype *A.B*, $^3/_{16}$ have phenotype *A.b*, $^3/_{16}$ have phenotype *a.B*, and $^1/_{16}$ have phenotype *a.b*. The reasons for these frequencies of phenotypes is explored in a paper by Dudley (1973a).

In classical experiments, such as those first performed by Mendel with round or wrinkled, yellow or green peas, and often repeated since with both plant and animal species, the four phenotypes occur amongst the offspring with a frequency of 9 to 3 to 3 to 1, and once again the unavoidable conclusion is that each gamete types of each parent occur with equal frequency. Since there are four gamete types produced by each of these parents, each type has a frequency of $^1/_4$. Since also alleles *A* and *a*, and *B* and *b*, must each been equally frequent (monohybrid inheritance) then it follows that half of the gametes which have allele *A* also have allele *B*, and the other half have *b*. The same applies to those which have allele *a*. Thus in forming the gametes, the alleles of the two genes must behave in the manner shown in the tree diagram



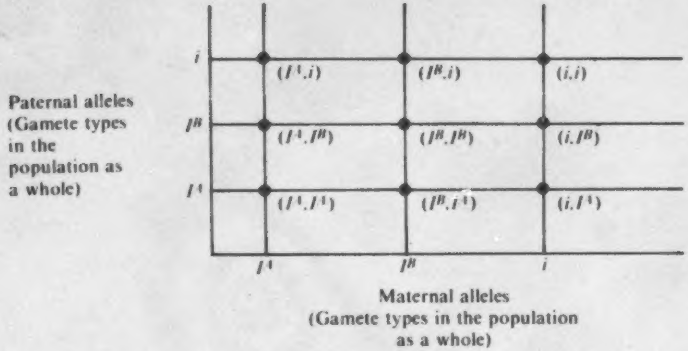
(Frequencies in brackets)

The genotypes of possible offspring of parents of the different experiments can be represented in this fashion, and this mode of presentation becomes useful also when a study is being made of those special cases of inheritance, such as multiple alleles, sex linkage and autosomal linkage that are part of many biology courses at secondary school.

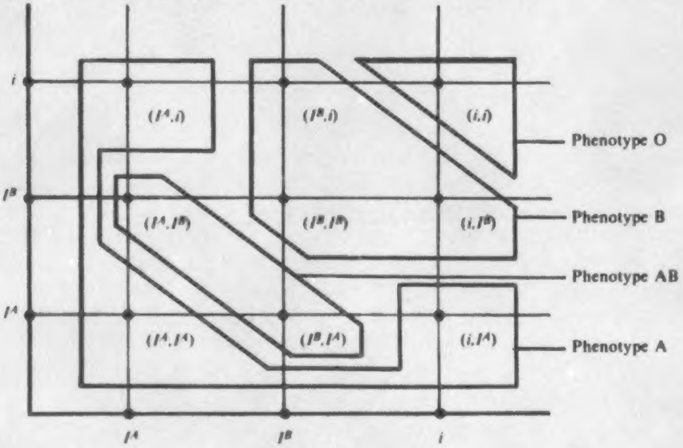
The genetics of multiple alleles

A multiple allele system is one in which an allele (e.g. *A*) has *more than* two expressions, say, *A,a,α* instead of just the more usual two expressions, *A* and *a*. Without doubt the best known example is that of the ABO system of human blood groups, though others exist, of which the system that includes the condition of rabbit's fur known as chinchilla is but one.

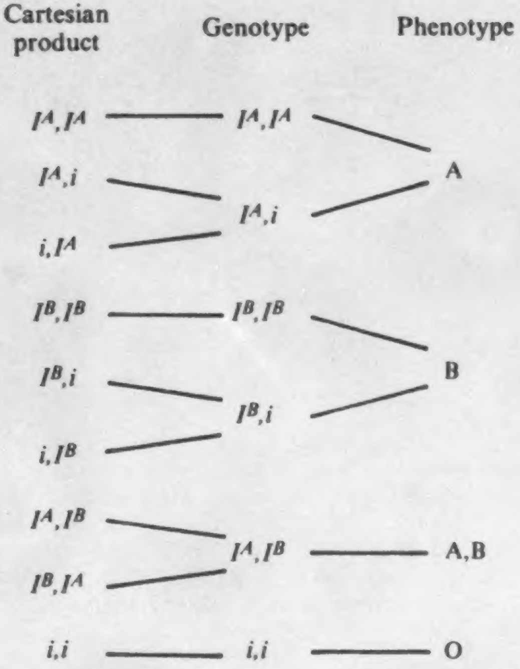
With regard to the ABO blood groups in humans, the maternal and paternal gametes of a population as a whole must be elements of the set $S = \left\{ I^A, I^B, i \right\}$. Each gamete carries only a single allele for this character. Under such a system all possible genotypes of the offspring that can arise are shown in the following lattice in which the different possibilities are each identified by a co-ordinate. The lattice represents the Cartesian product $S \times S = \left\{ I^A, I^B, i \right\} \times \left\{ I^A, I^B, i \right\}$.


Paternal alleles (Gamete types in the population as a whole)
i
I^B
I^A
(I^A,i)
(I^B,i)
(i,i)
(I^A,I^B)
(I^B,I^B)
(i,I^B)
(I^A,I^A)
(I^B,I^A)
(i,I^A)
I^A
I^B
i
Maternal alleles (Gamete types in the population as a whole)

Although there are 9 possible elements of the Cartesian product $S \times S$, so far as genetics is concerned there are only six different genotypes because some of the 9 are equivalent (for example (I^A,i) and (i,I^A) and also (I^A,I^B) and (I^B,I^A)). These six different genotypes result in there being four different blood groups because alleles I^A and I^B are expressed in the phenotype whenever they are present (they are co-dominant) but i is expressed only when I^A and I^B are absent and two alleles i are present instead. The four phenotypes (four blood groups) are A, B, AB and O and are shown below.


i
I^B
I^A
(I^A,i)
(I^B,i)
(i,i)
Phenotype O
(I^A,I^B)
(I^B,I^B)
(i,I^B)
Phenotype B
Phenotype AB
(I^A,I^A)
(I^B,I^A)
(i,I^A)
Phenotype A
I^A
I^B
i

The relationship between Cartesian product, genotype and blood type (phenotype) can also be portrayed as a mapping.



Knowledge of the inheritance of blood groups is used in cases of disputed parenthood. Once the blood group of the child and the mother are known then the necessary blood group of the father can be deduced. This is used in law as a means of *eliminating* from suspicion a man accused of being the father of that child. It cannot be used to prove who is the father, because it can only show possible groups of the father. If the accused has a different blood type from a list of those shown to be possible, then the case against the accused should be dismissed. If the accused has a blood type the same as that shown on the list, it still does not necessarily mean he is the father.

The complete set of possibilities is given in Table 2, which also illustrates a mapping of $S \times S \rightarrow S$, where S is the set of ABO blood groups. The symbol ϕ means that, in this instance, the case cannot arise.

Table 2

Identification of the blood groups of the father in disputed paternal cases (entries in the table give the possible blood groups of the father).

		ABO group of the child			
		O	A	B	AB
ABO group of the mother	O	A B O	A AB	B AB	ϕ
	A	A B O	A B O AB	B AB	B AB
	B	A B O	A AB	A B O AB	A AB
	AB	ϕ	A B O AB	A B O AB	A B AB

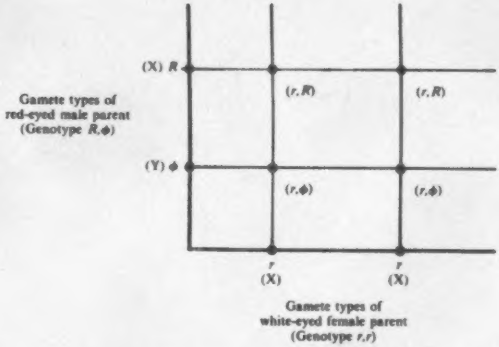
Sex linked inheritance

An allele is part of a chromosome. Two chromosomes, the X and the Y, are concerned with determining the sex of the individual. In most animals, including man, an individual with one X and one Y chromosome in each cell of the body is male, while one with two X chromosomes is female. However, these sex chromosomes also carry alleles of genes which are involved in determining other features of the organism. For instance, red-green colour deficiency (though not colour blindness) and haemophilia, are abnormalities in humans, the alleles for which occur on the X chromosome. Other alleles, determining the more normal conditions of sight and blood clotting exist, and, since it is these which are expressed in the phenotype whenever they are present in the genotype, are dominant to the alleles causing the defects mentioned. Alleles for such sex linked characteristics, whether dominant or recessive, are on that part of the X chromosome for which there is no counterpart on the Y chromosome, on the so-called differential part of the X chromosome; their inheritance, therefore, shows an unusual pattern, different from the classical Mendelian ratios associated with alleles on the "ordinary" chromosomes – the autosomes.

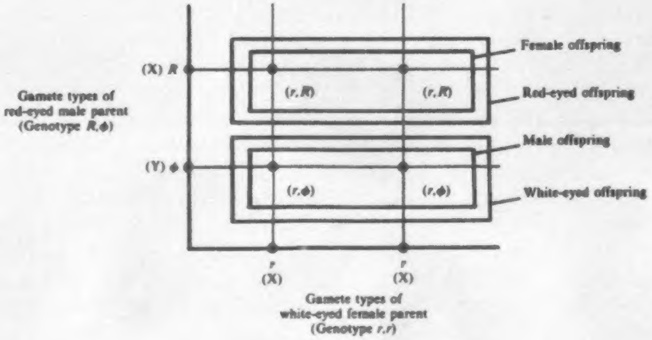
However, their inheritance is no more than a special case of monohybrid inheritance involving the empty set, ϕ, as can be illustrated with the traditional example of the eye colour of the fruit fly, *Drosophila melanogaster*, where white eye colour (r) is recessive to red eye colour (R).

A. A cross between a red-eyed male (R,ϕ) and a white-eyed female (r,r)

The male is X,Y and, at least as far as the genes on the differential part of the X chromosome are concerned, the corresponding part of the Y chromosome consists of an absent gene. The Cartesian product of such a cross (red-eyed male, white-eyed female) can be represented as follows


(X) R
(Y) ϕ
Gamete types of red-eyed male parent (Genotype R,ϕ)
(r,R)
(r,R)
(r,ϕ)
(r,ϕ)
r (X)
r (X)
Gamete types of white-eyed female parent (Genotype r,r)

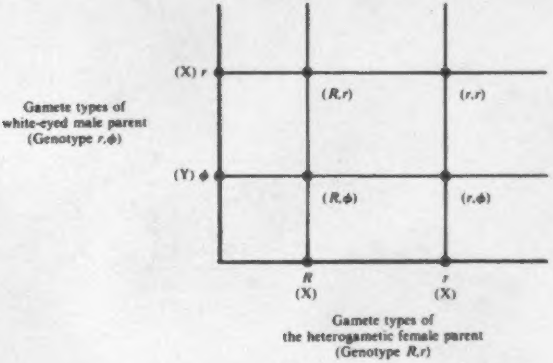
Ordered pairs that include the absent gene will be male, and since the only male genotypes in this cross are (r,ϕ) they all will have white eyes. Products without the absent gene will be female, and all such arising from this cross will have the genotype (r,R) and so will have red eyes. The offspring resulting from this cross will be such that all the males have white eyes and all the females red eyes.


(X) R
(Y) ϕ
Gamete types of red-eyed male parent (Genotype R,ϕ)
(r,R)
(r,R)
(r,ϕ)
(r,ϕ)
Female offspring
Red-eyed offspring
Male offspring
White-eyed offspring
r (X)
r (X)
Gamete types of white-eyed female parent (Genotype r,r)

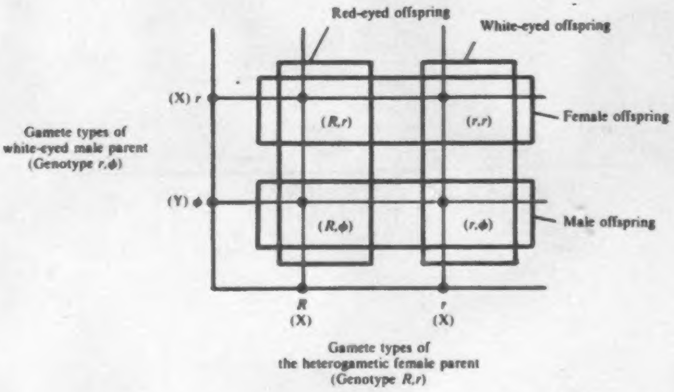
Incidentally, since each gamete type in each parent is equally frequent (and hence has a frequency of ½), then each of the four products will have a frequency of ¼ (½ × ½), when the phenotypes of the offspring will also be equally frequent, half being red-eyed female, half being white-eyed males.

B. A cross between a white-eyed male (r,ϕ) and a heterogametic female (R,r)

If the offspring of the first cross are bred to maturity and mated amongst themselves, then the types of offspring possible from such a cross are shown in the lattice:


(X) r
(Y) ϕ
Gamete types of
white-eyed male parent
(Genotype r,ϕ)
(R,r)
(r,r)
(R,ϕ)
(r,ϕ)
R
(X)
r
(X)
Gamete types of
the heterogametic female parent
(Genotype R,r)

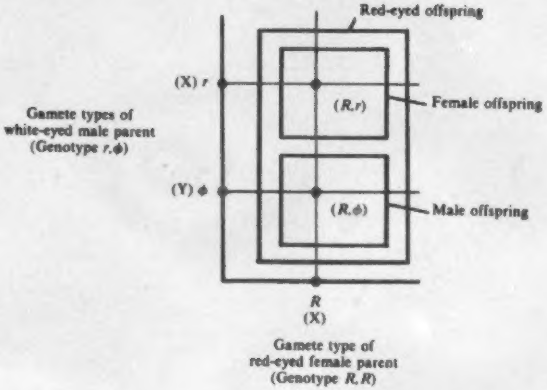
Again those products with the absent gene will be male, and those without the absent gene will be female. Those with the dominant allele, *R*, present will be red-eyed while those with only the recessive allele present, *r*, will be white-eyed. The frequency with which *each* product can be expected to arise in a large population of offspring is ¼, for reasons already given. The characters will be distributed amongst the offspring in the following fashion


Red-eyed offspring
White-eyed offspring
Female offspring
Male offspring
(X) r
(Y) ϕ
Gamete types of
white-eyed male parent
(Genotype r,ϕ)
(R,r)
(r,r)
(R,ϕ)
(r,ϕ)
R
(X)
r
(X)
Gamete types of
the heterogametic female parent
(Genotype R,r)

Half the male offspring will have red eyes, half white eyes, and the same will occur amongst the female offspring, and males and females themselves will be equally frequent.

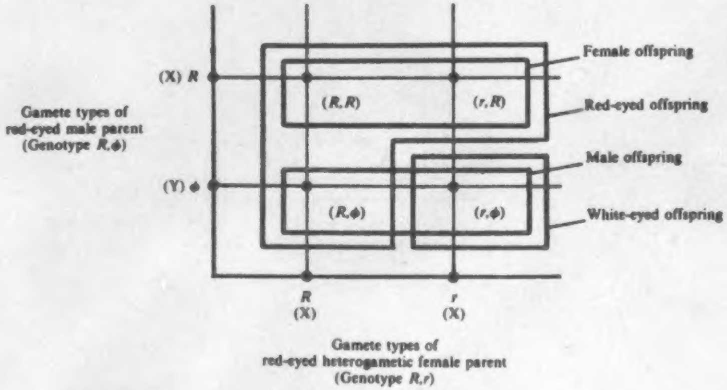
C. A cross between a white-eyed male (r,ϕ) and a pure bred red-eyed female (R,R)

The offspring of this cross will all have red eyes. On this occasion, with the female homogametic, it is possible to unite the sets of female gametes to give


Red-eyed offspring
(X) r
(R,r)
Female offspring
Gamete types of
white-eyed male parent
(Genotype r,ϕ)
(Y) ϕ
(R,ϕ)
Male offspring
R
(X)
Gamete type of
red-eyed female parent
(Genotype R,R)

D. A cross between a red-eyed male (R,ϕ) and a heterogametic red-eyed female (R,r)

In this instance the only white-eyed individuals amongst the possible offspring are male (r,ϕ); all the female offspring and half of the male offspring will have red eyes. This is seen to be the case from the following lattice.


Female offspring
(X) R
(R,R)
(r,R)
Red-eyed offspring
Gamete types of
red-eyed male parent
(Genotype R,ϕ)
Male offspring
(Y) ϕ
(R,ϕ)
(r,ϕ)
White-eyed offspring
R
(X)
r
(X)
Gamete types of
red-eyed heterogametic female parent
(Genotype R,r)

A knowledge of both sex linkage, and monohybrid inheritance are important in plant breeding, animal breeding, and human genetic counselling. Concerning the latter, many, but by no means all, of the congenital errors in humans are controlled by the recessive allele of a gene. (1) Suppose a wife has a brother who himself has haemophilia. The husband had a brother who died of haemophilia when young. What is the likelihood of the children of this marriage having *a*) a son with haemophilia, *b*) a daughter with haemophilia, *c*) a son who is a carrier for haemophilia (ie, is heterogametic for this character), *d*) a daughter who is a carrier for haemophilia? Suppose they have three daughters. Of *their* total alleles for this character, what proportion are likely to be haemophiliac alleles? These questions and those which follow, can be answered from the information provided so far.

(2) Phenylketonuria is a very rare disease which causes the central nervous system in humans to develop abnormally, usually resulting in a severely mentally retarded child. Suppose two normal parents produce an offspring that has phenylketonuria. What are the chances of a second child also having phenylketonuria, and what are the chances of the second child being *a*) a carrier for this condition, *b*) free, genetically, from this condition. Phenylketonuria is caused by a single recessive allele which is carried on an autosomal chromosome.

(3) In one family tree, the wife's brother has red-green colour deficiency, and so does the husband. The respective mothers, of the wife and of the husband have normal colour vision, and so do the respective fathers. From which side in the grandparent generation is this defect inherited? The wife and husband have two sons. What are the chance that *a*) either *b*) both will have red-green colour deficiency? What could be the genotype of any daughter by this marriage? The red-green colour deficiency is very marked on the wife's side of the family. On the husband's it is very slight. If you find the sons and/or the daughters can/will have colour deficiency or will be carriers for it, is it the marked deficiency they would each have or carry, or is it the slight defect?

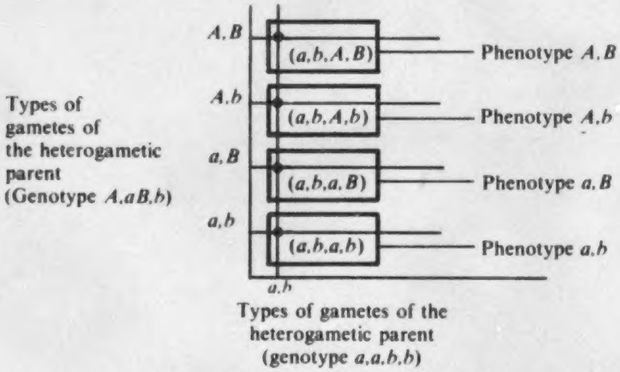
Concern amongst parents and prospective parents for the genetic well being of their offspring is very real and they can be as concerned about producing a carrier of some inherited defect as they are about producing a child with such defects. At the same time, a great deal of needless anxiety can be avoided if one knows the basic rules of the inheritance of single genes, whether they be part of the differential portion of the X chromosome (sex linked), or part of the autosomes ("normal" chromosomes).

Linkage

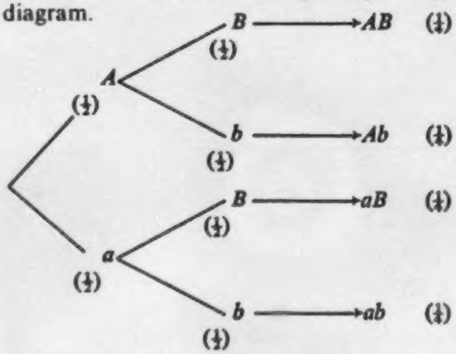
There is one particular kind of cross, involving dihybrid inheritance, which is of special interest to biologists because it enables one to find out the position and the sequence of genes on the chromosomes, and even the relative (chromosomal) distance between them. The genetic cross from which such information flows is one in which one parent has the genotype *A,a,B,b* (and is said to be heterogametic for both genes) and the other has the genotype a,a,b,b (and is said to be homogametic and recessive for both genes). It can be represented by the notation $A,a,B,b \times a,a,b,b$, and it follows that the types of gametes that can be formed in the first parent are (A,B), (A,b), (a,B), and (a,b) while those of the second parent are all of the type (a,b).

This kind of cross, and the types of possible offspring that can arise from it are shown in the Cartesian lattice below and its special interest lies in the fact that the phenotype of each offspring is governed by the alleles inherited from the heterogametic parent.

Products of the cross $A,a,B,b \times a,a,b,b$


A,B
A,b
a,B
a,b
Types of gametes of the heterogametic parent (Genotype A,aB,b)
(a,b,A,B)
(a,b,A,b)
(a,b,a,B)
(a,b,a,b)
Phenotype A,B
Phenotype A,b
Phenotype a,B
Phenotype a,b
a,b
Types of gametes of the heterogametic parent (genotype a,a,b,b)

Because the phenotype of each offspring is governed by the alleles contributed by the heterogametic parent, then the frequency with which each phenotype occurs is governed by the frequency of the corresponding gamete in the heterogametic parent. When all four phenotypes occur with equal frequency in a large population of offspring it means the four types of gamete must occur with equal frequency in the heterogametic parent, and the alleles of each gene must have combined at random in forming the gametes, as indicated in the following tree diagram.


B
AB
(¼)
(½)
A
(½)
b
Ab
(¼)
(½)
B
aB
(¼)
(½)
a
(½)
b
ab
(¼)
(½)

(Frequencies in brackets)

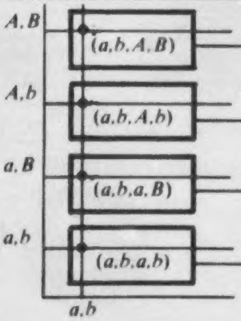
This result, with the four phenotypes equally frequent amongst the offspring of a cross of the type *A,a,B,b* × *a,a,b,b*, is said to be one involving classical mendelian inheritance, and its ratio of 1 to 1 to 1 to 1 (in biology written 1 : 1 : 1 : 1) is the classical ratio associated with it.

However, since the phenotypes that appear in the offspring map on to the gametes produced by the one parent, then it can, and is, also used to find out what is the genotype of individuals whose genotype cannot be deduced from its appearance (its phenotype). Such a cross is a Test cross. If the test individual has the phenotype AB, then one cannot tell from it which other alleles are present in its genetic make-up (in its genotype). Should this information be needed, the individual is crossed with another, whose genotype is *a,a,b,b*. One expects individuals amongst the offspring to have the phenotype AB, and if that is the only kind to appear in a large population of offspring, then the test parent can only have been producing the one, *AB*, type of gamete, and so its genotype must have been *A,A,B,B*. On the other hand if for instance two genotypes appear, *A,B* and *a,B* then, by the same reasoning the genotype of the test parent must be *A,a,B,B*. (Incidentally, what would be the frequency with which these three phenotypes occur in the offspring population?).

Sometimes, even with a heterogametic parent, the cross *A,a,B,b* × *a,a,b,b* consistently fails to yield the classical 1 : 1 : 1 : 1 ratio, and the first of these exceptions (Morgan, 1910) is considered in Table 3.

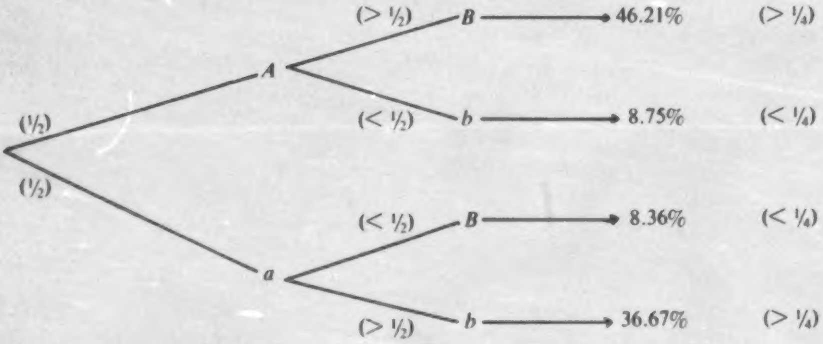
Table 3

Results obtained by Morgan (1910), on the fruit fly *Drosphila melanogaster*, an exception to those expected of classical Mendelian inheritance.



Cross (*A,a,B,b* × *a,a,b,b*)	Phenotypes of the offspring	Numbers obtained	Percentage incidence of phenotypes	Frequency of gamete types in heterogametic parent
(*a,b,A,B*)	*A,B*	586	46.21	> ¼
(*a,b,A,b*)	*A,b*	111	8.75	< ¼
(*a,b,a,B*)	*a,B*	106	8.36	< ¼
(*a,b,a,b*)	*a,b*	465	36.67	> ¼

Since the total frequency (P) of the alleles is such that $P(A) \simeq P(a)$, (54.96% (A), 45.23% (a)), and $P(B) \simeq P(b)$ (54.57% (B), 45.42% (b)), then in forming the gametes the alleles must have behaved in the following fashion



(Frequencies in brackets)

Thus (A and B) and (a and b) respectively are not inherited independently, but tend to stay together, so that the alleles in the heterogametic parent must behave in much the same way. These, and other exhaustive tests, caused Morgan to propose they remain associated (linked) by being on the same chromosome, and this was confirmed in subsequent experiments. If, however, they were completely linked there would have been only two classes (sets) of phenotype amongst the offspring generation. In fact, there are four, and this causes the phenomenon to be called 'incomplete linkage'. The existence of the two least frequent classes is due to a phenomenon whereby the chromosome sometimes breaks and, in rejoining, parts are exchanged (recombined) with other chromosomes.

If it is assumed that the breaks occur at random throughout the length of the chromosome, the incidence of these two least frequent classes can be used as a measure of the length of chromosome which lies between the two genes in question. By studying the inheritance of three genes simultaneously (a three point cross) the linear sequence and their relative distances can be established which form the basis of a map. It is in this way that chromosome maps are built up, at least for animals and for plants.

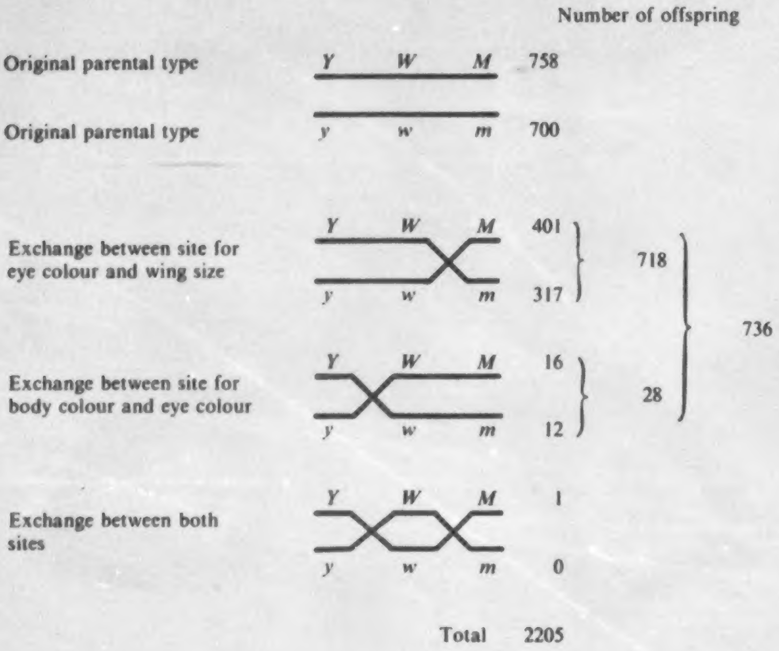
By way of illustration, one three point is considered in Table 4.

Table 4

Data from a three point cross performed by Morgan (1911). The cross is (*Y,y, W,w, M,m* × *y,y,w,w,m,m*).

Cross	Phenotypes of the offspring	Appearance Body colour	 Eye colour	 Wing size	Numbers		
Y,W,M (*y,Y,w,W,m,M*)	*Y,W,M*	Grey (Y)	Red (W)	Large (M)	758		
y,w,m (*y,y,w,w,m,m*)	*y,w,m*	Yellow (y)	White (w)	Miniature (m)	700		
Y,W,m (*y,Y,w,W,m,m*)	*Y,W,m*	Grey	Red	Miniature	317	718	736
y,w,M (*y,y,w,w,m,M*)	*y,w,M*	Yellow	White	Large	401		
Y,w,m (*y,Y,w,w,m,m*)	*Y,w,m*	Grey	White	Miniature	12	28	
y,W,M (*y,y,w,W,m,M*)	*y,W,M*	Yellow	Red	Large	16		
Y,w,M (*y,Y,w,w,m,M*)	*Y,w,M*	Grey	White	Large	1		
y,W,m (*y,y,w,W,m,m*)	*y,W,m*	Yellow	Red	Miniature	0		
y,w,m				Total	2205		

Since this cross deals with three linked genes, it is helpful to show them as such, and to show where each exchange of chromosomes is possible, along with the corresponding data:


Number of offspring
Original parental type
Y W M
758
Original parental type
y w m
700
Exchange between site for eye colour and wing size
Y W M
401
y w m
317
718
736
Exchange between site for body colour and eye colour
Y W M
16
y w m
12
28
Exchange between both sites
Y W M
1
y w m
0
Total 2205

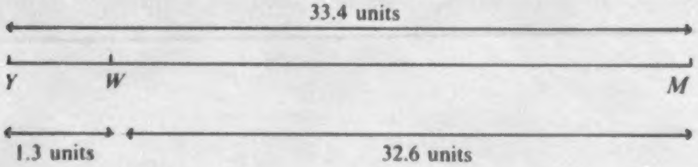
Recombinations and their incidence:

of *Y* and *W* $\quad 12 + 16 + 1 = \frac{29}{2205} \times 100 = 1.3\%$

of *W* and *M* $\quad 317 + 401 + 1 = \frac{719}{2205} \times 100 = 32.6\%$

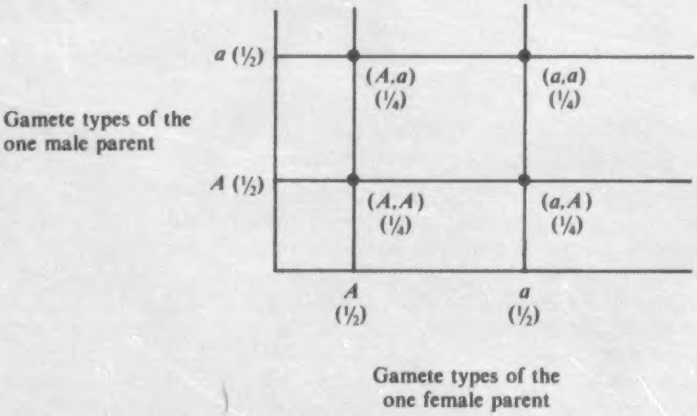
of *Y* and *M* $\quad 28 + 718 = \frac{736}{2205} \times 100 = 33.4\%$

It follows that the greatest chromosomal distance is between the site for body colour and the site for wing size, with the site for eye colour between the two. This gives the linear sequence of these three genes. The percentage recombinations gives the relative distances and so completes the chromosome map for these three genes which appears thus


33.4 units
Y
W
M
1.3 units
32.6 units

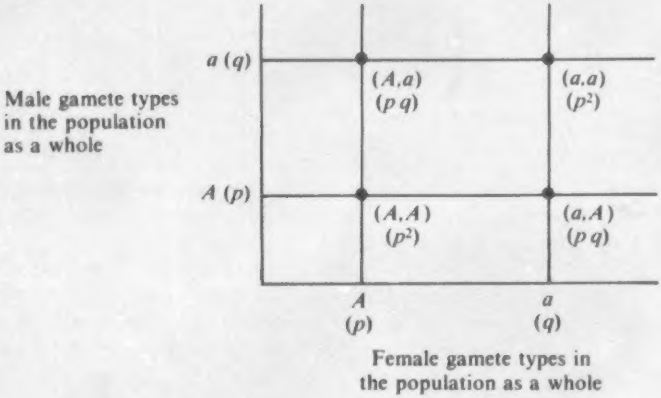
Population genetics

The study of the genetics of whole populations arises from the fact that monohybrid inheritance can be seen to be a particular instance of a more general case. In the particular instance the population is reduced to its limit, with one male and one female parent. In such circumstances, and with both parents of genotype *A*,*a*, there are four types of product, and these are (*A*,*A*), (*A*,*a*), (*a*,*A*) and (*a*,*a*). Moreover ¾ of the offspring will have the phenotype *A*, and ¼ phenotype *a* because each product arises with a frequency of ¼ since in each parent the alleles are equally distributed amongst the gametes, $^P(A) = {}^P(a) = \frac{1}{2}$ where *P* represents the frequency in accordance with the lattice


a (½)
A (½)
Gamete types of the one male parent
(A,a) (¼)
(a,a) (¼)
(A,A) (¼)
(a,A) (¼)
A (½)
a (½)
Gamete types of the one female parent

(Figures in brackets represent frequencies)

If, however, a study is made of the population as a whole, rather than of the offspring of only two individual parents, the products are as above, but the relative frequencies amongst the gametes of alleles *A* and *a* are unknown. If amongst the male gametes in the population the frequency of allele *A* is *p*, and the frequency of allele *a* is *q*, and assuming the distribution of these alleles amongst the female gametes is similar, then the result amongst the offspring, shown below, is immediately recognizable by biologists as the Hardy-Weinberg law, and by mathematicians as the expansion of the binomial $(p + q)^2$.


a (q)
Male gamete types in the population as a whole
A (p)
(A,a) (p q)
(a,a) (p²)
(A,A) (p²)
(a,A) (p q)
A (p)
a (q)
Female gamete types in the population as a whole

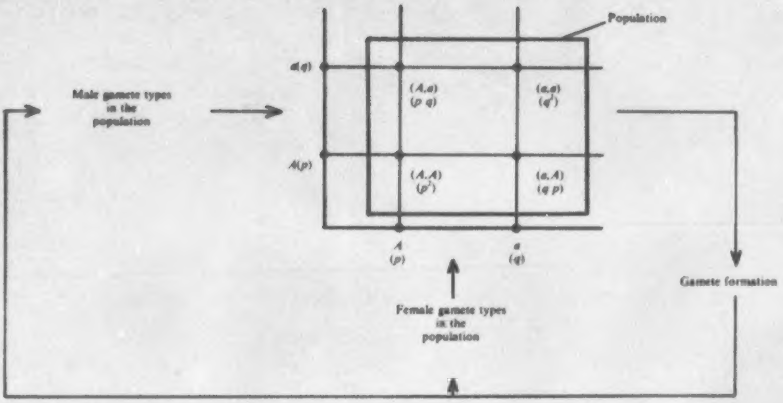
(Frequencies are shown in terms of *p* and *q*)

The frequency of (A,A) is p^2

The frequency of (A,a) is $2pq$

The frequency of (a,a) is $q^2 = (1 - p)^2$ since $p + q = 1$.

The studies of the patterns of inheritance undertaken by Mendel constitute a special instance of this general case, namely one in which $p = q = \frac{1}{2}$. Moreover, when the population of offspring themselves form gametes, the distribution of the alleles *A* and *a* amongst them remains unaffected, demonstrating one of the principles of the Hardy-Weinberg law which is that, given the parameters of the model, the genetic composition of the population *cannot* change and the population cannot evolve. That this must be the case can be represented in a single diagram (Figure 3), in which the Hardy-Weinberg law is summarised. Interestingly, it can be proved that the parameters of the model *cannot* be fulfilled in their entirety and hence the genetic composition of a population *must* always be changing.



The Hardy-Weinberg law, also a model, is used to gain insights into the genetic composition of populations, both human and non-human, and this information forms the bases of many decisions.

Suppose a farmer has a flock of chickens and is concerned by the occasional appearance of an undesirable phenotype, which occurs in 1% of his flock. He has found by breeding experiments that such individuals have a double recessive genotype (a,a).

The farmer needs to remove the undesirable allele a from the flock for it to have any economic value and has two choices. He can kill the individuals which have this recessive phenotype and progressively purify his flock over a number of generations, or he can create a new pure flock by selective breeding from individuals of genotype A,A within his present flock. Which method will give him the quickest result?

He wishes to maintain a flock of 1000 chickens.

Answer: The incidence of carriers (genotype A,a) needs to be calculated

The incidence of genotypes a,a in the population is $1\% = 0.01$. According to the Hardy-Weinberg model, the incidence of genotype a,a is $(1 - p)^2$

when $(1 - p)^2 = 0.01$

$$1 - p = \sqrt{0.01} = 0.1$$

$$p = 1 = 0.1 = 0.9$$

Frequency of genotype $A,A = p^2 = (0.9)^2 = 0.81$ or 81%

Frequency of genotype $a,a = 0.01$ or 1%

Frequency of genotype $A,a = 1.00 - (0.81 + 0.01) =$
$1.00 - 0.82 = 0.18$ or 18%

In a flock of 1000 chickens the number of poultry showing the undesirable phenotype is 1% of 1000, which is 10, while the number of carriers is 18% of 1000, which is 180. Killing the ten will leave unaffected the 180 carriers. It is unlikely that the allele a will be removed from the stock rapidly by removing only those poultry with the phenotype a. The farmer would need to breed from poultry of the genotype A,A and he would identify these by means of the test cross.

In Britain, it is estimated that about 1 child in 1500 is born with a condition known as *cystic fibrosis*. This is a condition caused by the genotype a,a, and is often fatal before the child reaches adolescence. The Hardy-Weinberg model enables one to estimate the proportion of adults that are carriers.

The incidence of cystic fibrosis is 1 in 1500 $= \frac{1}{1500} = 0.0007$

According to the Hardy-Weinberg model

Incidence of genotype $a,a = (1 - p)^2 = 0.0007$

when $1 - p = \sqrt{0.0007} = 0.02569$

$p = 1 - 0.02569 = 0.9744$

Frequency in the population of genotype A,A is $p^2 = (0.9744)^2 = 0.9495$

Frequency in the population of genotype a,a is 0.0007

Frequency in the population of genotype A,a (carriers of cystic fibrosis) is $1 - (0.9744 + 0.0007) = 0.05$ (approx.)

According to the model approximately 5% or 5 people in every 100 (1 in 20) are carriers for this congenital error in Britain.

The incidence of other recessive congenital errors in human populations are considered to be as follows

Phenylketonuria	1 in 25,000 births
Albino	1 in 20,000 births
Amaurotic idiocy	1 in 200,000 births

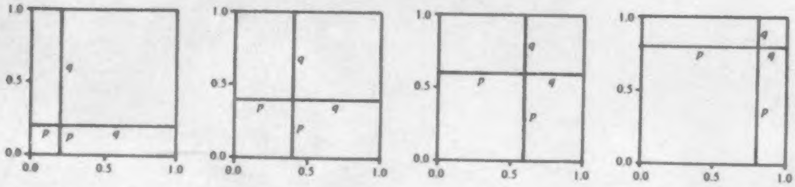
What is the incidence in the human population of the carriers of *each* of these errors? Express the answer as the number per 100 and as the number per 1000.

What is the incidence in the human population of carriers for *two* or for all *three* of these errors. What is the incidence of individuals who are carriers for phenylketonuria *or* albino *or* amaurotic idiocy?

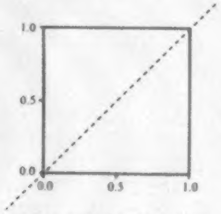
The diagram of the genetic composition can be presented as a square (as well as a Cartesian product) where the areas of the square portray the frequency of the different genotypes and the two axes are linear scales, each reading from 0.0 to 1.0.

90

The following represent the range of squares arising

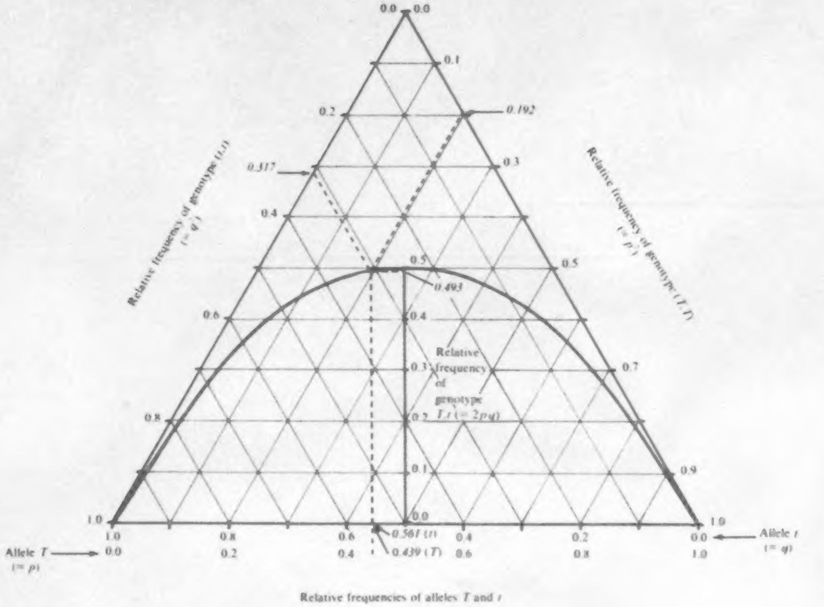


These figures show that the squares are symmetrical along the one diagonal because it is along this diagonal that the lines always intersect. One half of the square is a reflection of the other:



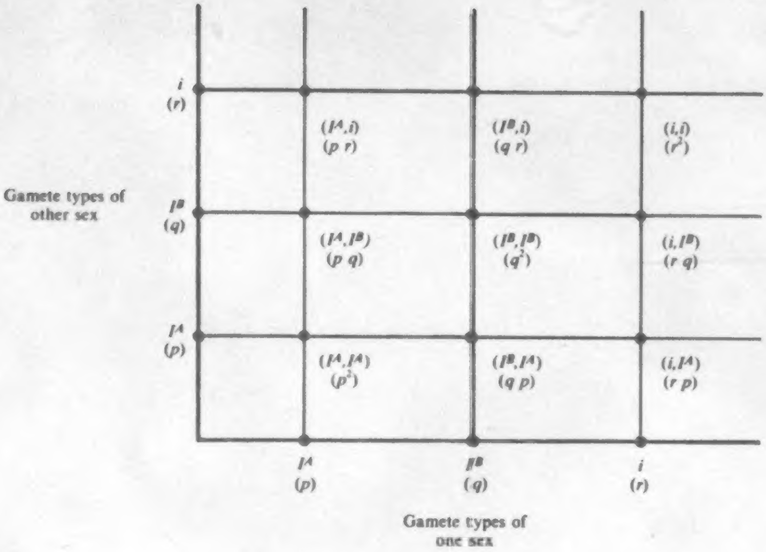
Axis of symmetry of the figures above

It is possible, therefore to represent the genetics of a population by means of a triangle in place of a square, but it is convenient to use a symmetrical (isosceles) triangle rather than a right angle triangle, and to this end isometric graph paper is especially useful. Each side becomes a scale line representing, respectively, the frequency (*P*) of genotype *A,A*, the frequency (*P*) of genotype *A,a*, and the frequency (*P*) of genotype *a,a*. If also the line showing the composition of a population according to the Hardy-Weinberg model is drawn on the graph, then it is a simple matter to read the frequency of the other two genotypes as soon as the one is known. Usually it is the frequency of genotype *a,a* that becomes known first because these all have the recessive phenotype and are readily recognised in a population, so one reads the graph from the axis *P*(*a,a*), and from that reads the values not only of *P*(*A,A*) and *P*(*A,a*) – the carriers – but also the frequency of allele *A* and allele *a* in the population as a whole. All this information is valuable in the study of population genetics and especially in the changes in the genetic composition of that population over a period of time.



The case shown on the graph is that investigated by Snyder (1931) in America, who investigated the incidence of those who can, and those who cannot, taste a solution of phenyl-thiocarbamide (PTC). Non-tasters have the recessive genotype *t,t* while tasters have the allele *T*, and may have the genotype either *T,t* or T,T. Snyder found 31.7% of his large sample to be non-tasters and so, from the Hardy-Weinberg model, 19.2% have genotype *T,T*, 49.3% have genotype *T,t* (i.e. are carriers of the non-taster allele), while the frequency of the allele *T* in the population as a whole is 0.439 (= p), and of the allele t is 0.561 (= q). While this can be calculated as shown earlier, it is generally more convenient to read it from the graph.

The use of the Cartesian product to whole populations is capable of further extension, which also serves to highlight its essential features. In the population genetics study of a multiple allele of the ABO system of blood groups in man, the population of the gametes of one sex is made up of more than two types, on this particular occasion, I^A, I^B, and i. Assuming the distribution amongst the gametes of each sex to be identical, then according to the model the frequency of the genotypes within the population are as follows:


i
(r)
(I^A,i)
(p r)
(I^B,i)
(q r)
(i,i)
(r^2)
Gamete types of other sex
I^B
(q)
(I^A,I^B)
(p q)
(I^B,I^B)
(q^2)
(i,I^B)
(r q)
I^A
(p)
(I^A,I^A)
(p^2)
(I^B,I^A)
(q p)
(i,I^A)
(r p)
I^A
(p)
I^B
(q)
i
(r)
Gamete types of one sex

where the frequencies in each sex of I^A gametes is (p), of I^B gametes is q, and of i gametes is r.

Genotypes amongst the offspring	Frequency
I^A,I^A	p^2
I^A,I^B	$2pq$
I^B,I^B	q^2
I^A,i	$2pr$
I^B,i	$2qr$
i,i	r^2
Sum of the frequencies $(p + q + r)^2$	

When two alleles are involved, the sum of the frequencies of the genotypes is $(p + q)^2$, when three are involved it is $(p + q + r)^2$. So, when n alleles are involved the sum of their frequencies can be expected to be $(p + q + r + \ldots\ldots + n)^2$, where $p, q, r, \ldots\ldots, n$, are the frequencies of the respective alleles in the population as a whole.

In Britain the mean frequencies for the four blood-groups are approximately 0.47 for O, 0.42 for A, 0.08 for B and 0.03 for AB. As for the population genetics of a two allele system, the ABO blood group system can usefully be represented as a latin square, but before that can be done the frequencies of the individual alleles needs to be calculated.

Phenotype	Genotype	Frequency
O	i,i	r^2
A	$I^A,I^A + I^A,i$	$p^2 + 2pr$
B	$I^B,I^B + I^B,i$	$q^2 + 2qr$
AB	I^A,I^B	$2pq$

Incidence of phenotype group O $= 0.47 = r^2$

$$r = \sqrt{0.47} = 0.686$$

Incidence of phenotype group A $= 0.42 = p^2 + 2pr$

Incidence of phenotype group O + A $= 0.47 + 0.42 = p^2 + 2pr + r^2 = (p + r)^2$

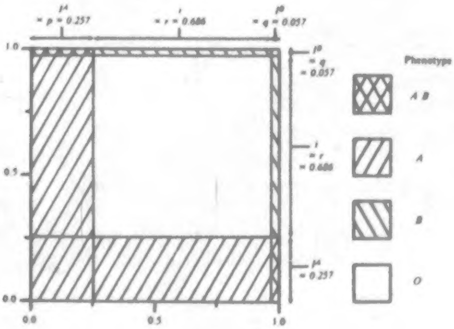
$$p + r = \sqrt{0.47 + 0.42} = \sqrt{0.89} = 0.943$$

$$p = 0.943 - 0.686 = 0.257$$

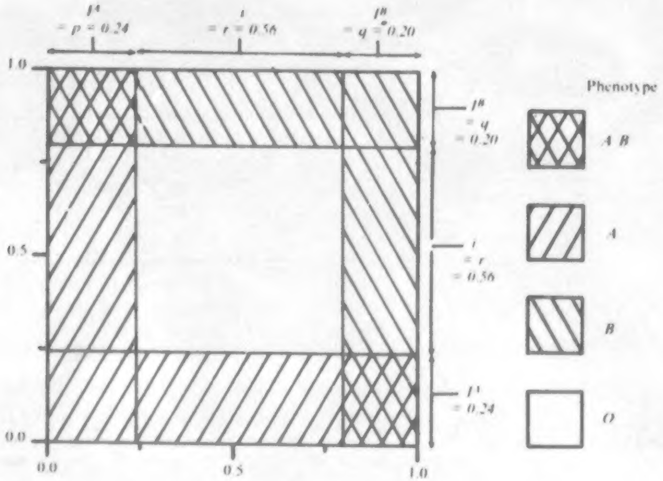
$$q = 1 - (p + r) = 1 - 0.943 = 0.057$$

So frequency of allele I^A is 0.257 (= p)
of allele I^B is 0.057 (= q)
of allele i is 0.686 (= r)

A latin square of representing the ABO blood group system in England has the following appearance, that for other countries will have a different appearance.



Amongst gypsies in Europe for instance the latin square has this appearance:



$I^A = p = 0.24$
$i = r = 0.56$
$I^B = q = 0.20$
1.0
0.5
0.0
0.0
0.5
1.0
$I^B = q = 0.20$
$i = r = 0.56$
$I^A = 0.24$
Phenotype
AB
A
B
O

The frequencies of the different alleles varies throughout the world, for reasons which are not immediately apparent.

As a point of general interest, the ABO blood type is one of the characters that has to be considered when a blood transfusion is contemplated, for some types react dangerously with others. A matrix summarising which types are compatible is shown below

		Donor			
		O	A	B	AB
Recipient	O	1	1	1	1
	A	0	1	0	1
	B	0	0	1	1
	AB	0	0	0	1

"1" denotes 'compatible', "0" denotes 'incompatible'.

It follows that group O is the 'universal recipient', at least so far as the ABO system is concerned.

References

Dudley, B., 1972, The mathematical basis of mendelian genetics, *Journal of Biological Education*, Vol 6, p. 165-173; also *International Journal on Mathematical Education in Science and Technology*, Vol 4, p. 193-204 (1973).

Dudley, B., 1973a, The mathematical basis of phenotype ratios. *Journal of Biological Education*, Vol 7, p. 25-28.

Dudley, B., 1973b, The mathematical basis of population genetics. *Journal of Biological Education*, Vol 7, p. 41-43; also *International Journal on Mathematical Education in Science and Technology*, Vol 4, p. 329-333 (1973).

Dudley, B., 1977a, Sorting dinosaurs: an opportunity to teach mathematics with biology. *International Journal on Mathematical Education in Science and Technology*, Vol 8, p. 417-443.

Dudley, B., 1977b, *Mathematical and Biological Interrelations*, Published by John Wiley & Sons, 319 pp.

Manton, I., and Sledge, W. A., 1955, Observations on the cytology and taxonomy of the pterydophyte flora of Ceylon. *Philosophical Transactions of the Royal Society* (B), Vol 238, p. 127-185.

Morgan, T.H., 1910, Sex limited inheritance in *Drosophila*, *Science*, N.Y., Vol 32, p. 120-122.

Morgan, T. H., 1911, An attempt to analyze the constitution of the chromosomes on the basis of sex-limited inheritance in *Drosophila*, *Journal of Experimental Zoology*, Vol 11, p. 365-414.

Rebman, K. R., 1967, 'Blood types as sets' in *some mathematical models in Biology* (Edited by Thrall, R. M., Mortimer, J. A., Rebman, K. R. and Baum, R. F.) University of Michigan.

Reid, D., and Booth, P., 1970, Biology for the Individual, Book 1. *Sorting Animals and Plants into Groups*, Heinemann, London, 1970.

Smith, A., 1975, *The human pedigree*, published by George Allen and Unwin, London, 327 pp.

Snyder, L. H., 1931, Inherited taste deficiency. *Science, N.Y.*, Vol 74, p. 151-152.

White, M. J. D., 1942, *The Chromosomes*, published by Methuen, London, 159 pp.

The development of school chemistry today is often characterised by new syllabuses which their instigators consider improvements. Some chemistry teachers are doubtful about new proposals, however, and choose not to adopt them, but many teachers, in different countries, seem happy to use ideas from modern syllabuses. Thus, in English speaking countries, the various British Nuffield and Schools Council schemes, or the American Chem Study and CBA programs, contain good news for the chemistry teacher.

The changes of emphasis and subsequent incorporation of new material may well be at the expense of traditional domains of knowledge and this has also happened in mathematics as in chemistry. Some of the ideas in new mathematics syllabuses may well be good news for the chemistry teacher. Often, however, the chemist who has been taught traditional mathematics is aware only of the deficiencies of students in such subject areas as algebraic manipulation and mechanical computation, without being aware of compensatory advantages of the new mathematics. For convenience we have divided this booklet into two distinct sections. The first deals with the 'traditional skills' needed in chemistry, and the second with *new* ideas from 'modern mathematics'.

Mathematics and Chemistry: The Classroom Interface

Gordon Woods
Monmouth School, Monmouth.

Section 1

Most of this section is specifically directed at those chemistry teachers whose *pupils* follow a modern mathematical course. It suggests different approaches and terminology in *traditional* areas of mathematical skills. There follows a section on graphs which is directed at both chemist and mathematician, irrespective of their mathematical background. In particular it makes suggestions for the plotting of experimental points. A problem list at the end of the booklet directed at mathematicians is a selection of chemical problems, mostly for pupils aged up to 16, chosen to illustrate different mathematical techniques. A chemistry teacher might find it convenient to hand this section, duly photocopied, to his mathematics colleagues. It could either be used in mathematics lessons to provide real examples or just to indicate to the colleague the numerate skills needed in chemistry but restricting their use to chemistry classes.

In general the chemists' requirements are for a small range of mathematical skills at a relatively low level but they must be at a high degree of competence, that is *simple* procedures *well done*.

Proportion

Possibly the most important mathematical skill needed in elementary chemistry is the carrying out of calculations involving proportions. From the equation

$$2\mathrm{Mg} + \mathrm{O_2} \rightarrow 2\mathrm{MgO}$$

the chemist deduces from relative atomic masses that 48g magnesium form 80g magnesium oxide. What mass of magnesium is needed to make 12g magnesium oxide? For some pupils the difficulty lies in the first step of obtaining the correct proportional expression $(48 \times \frac{12}{80})$, for others the problem is the second step of computing this expression. Many pupils are likely to describe the ratio of magnesium oxide in the new situation to the original situation $(\frac{12}{80})$ as the *scale factor*. They are less likely to be familiar with unitary method.

Note that the mass of oxygen required is obtained by multiplying the relative molecular mass of oxygen (32) by the *same scale factor*. (Some pupils use the term *enlargement factor*, but there is a potential mental blockage when the quantities are reduced, say halved, where the enlargement factor is ½.)

The second step of computing the expression can be done by the pupil himself in simpler cases. Harder examples have until recently needed a slide rule, but there is now a growing number of electronic calculators. Our experience is that a higher percentage of owners of calculators automatically bring them to non-mathematics classes than is the case with the slide-rule. Is this the convenience of their shape or their status? The calculator has boosted the confidence, even competence, of many pupils; furthermore it is particularly useful in real situations, as in practicals where it is not possible to organise the quantity of chemicals to ensure easy computation.

Re-arrangement of Formulae

How does a new mathematician pupil re-arrange the expressions $\frac{P_1V_1}{T_1} = \frac{P_2V_2}{T_2}$ to make V_1 the subject? He carries out a series of *identical operations to both sides*, thus

$$\frac{P_1V_1}{T_1} = \frac{P_2V_2}{T_2}$$

multiply both sides by T_1

$$P_1V_1 = \frac{P_2V_2T_1}{T_2}$$

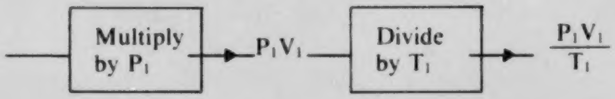
divide both sides by P_1

$$V_1 = \frac{P_2V_2T_1}{P_1T_2}$$

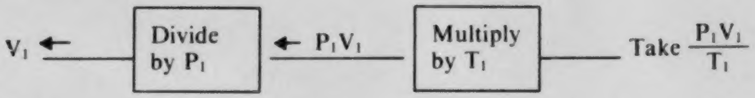
Note that he does *not* cross multiply.

Alternately the pupil may adopt a flow diagram approach in which he considers how the expression is built up. In each separate step a letter (or number) is fed into a "machine" which performs an 'operation' like multiplying, adding two, etc.

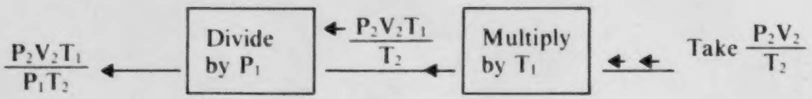
Thus take V_1



If one fed the answer $\left(\frac{P_1V_1}{T_1}\right)$ into the end of the machine and ran it backwards one will end up again with V_1: however when running backwards the operations must be made their respective *inverses*; 'multiply by' becoming 'divide by', 'add 2' becoming 'subtract 2' etc. Thus

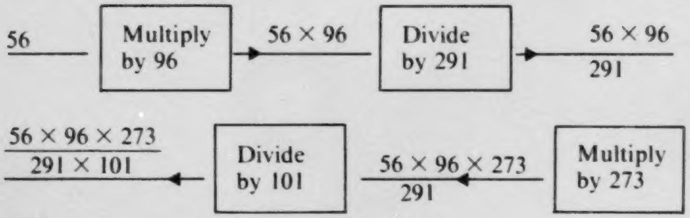


What happens if one feeds $\frac{P_2V_2}{T_2}$ into the right hand side?

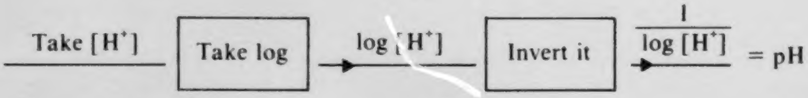

$\frac{P_2V_2T_1}{P_1T_2}$
Divide by P_1
$\frac{P_2V_2T_1}{T_2}$
Multiply by T_1
Take $\frac{P_2V_2}{T_2}$

This method of finding V_2 is probably unfamiliar to the chemistry teacher – it's probably no better than the traditional method he knows, but it *is* a method with which his new mathematics pupil is more familiar.

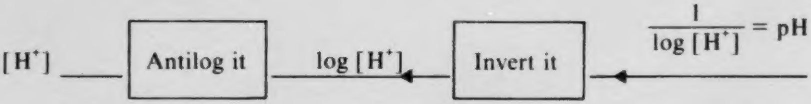
Many pupils are happier dealing with numbers than the symbols which represent them. Thus to find the volume at s.t.p. ($T = 273\,K$, $P = 101\,kN\,m^{-2}$) which corresponds to a volume of $56\,cm^3$, measured at a temperature of 291 K and $96\,kN\,m^{-2}$ pressure, a pupil knowing $\frac{P_1V_1}{T_1} = \frac{P_2V_2}{T_2}$ might write


56
Multiply by 96
56×96
Divide by 291
$\frac{56 \times 96}{291}$
$\frac{56 \times 96 \times 273}{291 \times 101}$
Divide by 101
$\frac{56 \times 96 \times 273}{291}$
Multiply by 273

More advanced pupils find it easier to obtain the pH for the hydrogen concentration than the reverse process. Furthermore some electronic calculators which have a log button do not have a single anti-log button. The flow diagram method with inverse function approach provides a simple method. Consider the process $pH = \frac{1}{\log[H^+]}$


Take $[H^+]$
Take log
$\log[H^+]$
Invert it
$\frac{1}{\log[H^+]}$ = pH

which when reversed becomes


$[H^+]$
Antilog it
$\log[H^+]$
Invert it
$\frac{1}{\log[H^+]}$ = pH

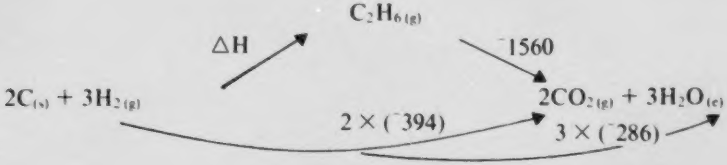
Notice 'invert it' is a self-inverse function. That is: its "opposite" is itself. (Compare rotating through 180°).

Plus and minus - correct usage

Modern mathematics draws a distinction between the *operations* add (plus) and subtract (minus) and the *signs* positive and negative. ($3 - 5$ is 3 minus 5, minus being the operation). Note that small plus and minus are used to indicate the sign (e.g. $^{+}20$°C, $^{-}10$°C). Chemistry teachers will lessen confusion in their pupils if they say negative 20°C for $^{-}20$°C and strictly Zn^{2+} should be described as a doubly positive ion, rather than two plus, though maybe this is being pedantically correct.

Example

A thermochemical calculation to find the enthalpy of formation of ethane can be represented diagrammatically.

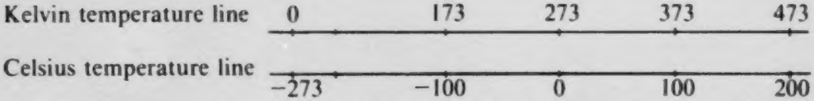

$C_2H_{6(g)}$
ΔH
$^{-}1560$
$2C_{(s)} + 3H_{2(g)}$
$2CO_{2(g)} + 3H_2O_{(e)}$
$2 \times (^{-}394)$
$3 \times (^{-}286)$

Apply Hess's Law one obtains $\Delta H + (^{-}1560) = 2 \times (^{-}394) + 3 \times (^{-}286)$ hence

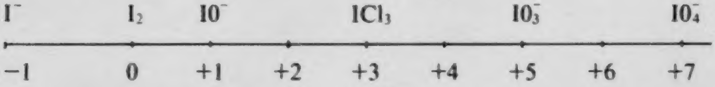
$$\Delta H = {}^{-}788 + {}^{-}858 - ({}^{-}1560) = {}^{-}788 + {}^{-}858 + 1560$$
$$= {}^{-}86$$

The number line

The number line is encountered at an elementary level. This idea can be applied to the interconversion of temperatures on different scales, thus


Kelvin temperature line
0
173
273
373
473
Celsius temperature line
−273
−100
0
100
200

At a more advanced level it can help in the calculation of electrode potentials of cells or more simply in what one might call 'oxidation number line diagrams'.


I^-
I_2
IO^-
ICl_3
IO_3^-
IO_4^-
−1
0
+1
+2
+3
+4
+5
+6
+7

Several small points

Pupils can be confused by trivial differences of motivation or approach, yet these often take time to discover by cooperation with one's mathematics colleagues — or try looking in a pupil's book.

(i) The equation for a straight line may be written $y = ax + b$ rather than $y = mx + c$.

(ii) The signs $\therefore$ and $\because$ have largely been replaced by $\Longrightarrow$ and $\Longleftrightarrow$ though with a different emphasis of meaning.

Thus $a \Longrightarrow b$ means if a then b.

and $a \Longleftrightarrow b$ means if a then b *and* if b then a.

A *tesselation* (literally a tiling) is a repeated 2 dimensional pattern. Thus a layer of graphite is a tesselation of hexagons.

Compared with traditional mathematics, a greater use is made of inequalities; nowadays pupils will be more familiar with $>$ and $<$.

Matrices are important in modern mathematics and yet they are of little help to the chemist. They can be used to balance chemical equations but this usage seems rather artificial.

Graphs

Chemistry teachers frequently use graphs as a method to display results. Sometimes a sufficient number of points stems from an experiment of an individual pupil or the demonstration of a teacher but on other occasions the experiment is such that each pupil only has sufficient time to obtain a single point, the graph then representing a collation of individual results within a form. Unlike the numerical data usually found in mathematical text-books, the scientists' practical data are subject to experimental error. Although it is more time-consuming it is easier to identify the individual bad result when all the results are plotted graphically, than when they are merely tabulated. The chemistry teacher is unwise to use graphical techniques with which his pupils are unfamiliar. Different points to be considered will now be mentioned with the aid of specimen data. The problem list contains most of the detailed data from which the graphs are plotted, along with further examples.

Choice and Labelling of Axes

Some mathematical colleagues assure me that 'plot temperature against time' implies that temperature is on the y axis and time on the x axis. However since not all my pupils to Ordinary level recognize this use of language, I prefer 'plot temperature (y axis) against time (x axis)'. Better pupils should be encouraged to choose for themselves, the convention being to plot the *independent variable* on the x axis. The idea of 'independent variable' is not so commonly used in mathematics today. Usually it refers in science to the variable which we measure first, usually in equal intervals. In mathematics the variables are (strictly speaking) interdependent!

It is now generally recognised, and stressed in the Metrication Board's recent publication, that graphs plot *numbers* and not quantities, so that one should encourage the labelling method of 'quantity measured/units in which it is measured', e.g. time/seconds, temperature/°C. Initially pupils can be instructed in this method as a routine, rather than understanding that the "/" converts the quantity measured to a dimensionless number.

Choice of Scale

Graphs should fill the paper available, with the slope of straight lines within 30° − 60° range with a simple scale. With normal graph paper with 10 small divisions to 1 large, it is advisable to choose 1 large division as 10, 5 or 2 units, with appropriate powers of ten, so that 1 unit is an exact number of small squares, respectively 1, 2 or 5. Some mathematics departments initially use 0.5 cm squares, *without subdivisions*. This is very uncluttered for display purposes but it is unwise for scientists to use subdivided graph papers before it has been used in mathematics.

The chemistry teacher can ease the choice of scale by counting the large squares on the graph paper and suggesting an appropriate range of readings. For example given 15 cm × 20 cm graph paper, with margins, an experiment may conveniently be carried out for 15 or 20, 30 or 40 minutes.

Should the zero be shown?

In general, chemists are interested in *changes* of mass, length, temperature, time, etc., so the zero is shown if it occurs, or nearly occurs, during the experiment; sometimes a zero and a broken line shows that the results are not continuous (see Graph 1 (b-d)). If a stop-watch is used to record the passage of time the reading would be directly plotted, even if an ordinary wrist-watch were used, the elapsed time rather than the time of day would be put on the graph.

One Nuffield experiment investigates the variation with elapsed time, of the mass of a flask containing hydrochloric acid and solid calcium carbonate. A loss of mass occurs because (carbon dioxide) gas is liberated. Some specimen results are

Time / second	0	30	60	90	120	180	240	300 [= ∞]
Total mass / g	107.54	106.88	106.56	106.36	106.28	106.19	106.16	106.14

How are these results best plotted? Graph 1a shows them directly plotted: very illuminating! One suggestion is to subtract a nominal 100 g which does focus attention on the relevant portion of the graph. However starting the graph at 106.14 g, the final mass of the flask and contents, the y axis now becomes the mass of unreacted gas (still as combined solid) Graph 1c. (Some modern top-pan chemical balances have a taring facility by which the zero can be adjusted to the mass at the start of the experiment, but this is more directly useful to the reverse situation when the mass is increasing to a maximum).

The ultimate purpose of this experiment is to investigate the effect of the lump size of the calcium carbonate. Thus the experiment is repeated in another similar flask with the 'same' amounts of the two chemicals, but with the calcium carbonate crushed. Because the mass of flasks and the chemicals will be slightly different, the results would be like this

Time / second	0	15	30	60	90	120	180 [= ∞]
Total mass / g	108.26	107.40	107.16	106.98	106.90	106.87	106.86

How do we plot them? Not with the same y axis, but we again subtract the final mass of the flask and contents, so that the scale is unaltered but the zero is displaced by a different amount. The results (Graph 1d) which can be superimposed on Graph 1c, show that the total amount of gas liberated is unchanged but at a greater rate when the lumps are smaller.

Marking points

My mathematical colleagues encourage their pupils to plot points with a + rather than a ⊙ or × or ▽ (These alternate methods are useful in print to distinguish between different sets of points where in class pupils should be encouraged to use different coloured *fine* biros or felt-tips). The + is recommended because it is made by first finding the values on the two axes, then moving along the line perpendicular to the axes, these two lines intersecting at the desired point. Whichever symbol is used it should be drawn large to indicate the experimental error.

A common experiment is to investigate the increase in mass, through gain of oxygen from the air, when magnesium is heated. The table below lists some results obtained by a competent junior form.

Mass magnesium / g	0.40	0.70	0.80	0.90	1.10	1.10	1.18	1.30	1.50
Mass oxygen / g	0.36	0.42	0.42	0.60	0.30	0.76	0.78	1.00	0.96

Graph 2a is obtained when these values are plotted as points, and Graph 2b when crosses are used. In this latter case each arm of the + is plotted to represent 0.1 g, i.e. an accuracy of ±0.1 g each direction. It is helpful to discuss the likely accuracy of the results to *roughly* decide how large to draw the + . A refinement which may be introduced for certain experiments for which one believes that one of the measured variables is more liable to inaccuracy, is to draw the + with arms of unequal length, thus + or +. For most pupils this refinement is unlikely to be other than an ideal to aim for.

Another advantage for large 'points' is that the best fit trace, whether it is straight or curved may be made to go through part of the 'point', when using the blackboard, the teacher can use the chalk *on its side* to draw a line, say 10 cm wide, covering the points. When using an overhead projector, a transparent plastic ruler, ideally of width representing the estimated spread of the points, is laid over the points and parallel lines are drawn along its edges, creating a *tunnel* to enclose the points. Pupils can follow a similar technique with a transparent set-square if their ruler is opaque. Of course the freehand chalk trace can be curved, the ruler line must be straight.

The burning of magnesium oxide is sometimes used as an experimental justification for the Law of Constant Composition, that all specimens of a given compound, magnesium oxide in this case, contain the same elements in the same proportion by mass.

$$\text{i.e.} \quad \frac{\text{mass (oxygen)}}{\text{mass (magnesium)}} \text{ is a constant}$$

The graph being a straight line allowing for experimental error, shows the constancy of the ratio, the slope of the line being the ratio.

Chemists take note! It is only correct to say y is proportional to x for a straight line *through the origin*, when y is plotted against x. A straight line *not* through the origin is described as "y is a linear function of x."

Average pupils often experience difficulty in obtaining satisfactory results first time; the chemistry teacher can improve the *appearance* of the results by plotting mass magnesium against *magnesium oxide*, overtly because these are the measured quantities but partly because the inclusion of magnesium within the mass of magnesium oxide imparts a gloss to the results.

We can write composition is constant $\Longrightarrow K = \dfrac{\text{mass (magnesium oxide)}}{\text{mass (magnesium)}}$

$$= \frac{\text{mass (magnesium)} + \text{mass (oxygen)}}{\text{mass (magnesium)}}$$

$$= 1 + \frac{\text{mass (oxygen)}}{\text{mass (magnesium)}}$$

Graphs 3a and 3b illustrate this point, they represent the *same* experimental results.

Gradients and Tangents

The chemistry students' first extensive use of gradients of a graph is in the investigation of reaction rates, in which the quantity of substance produced, or remaining, is plotted against time. It is important for time to be the x axis so that the gradient is the rate, as with the hydrochloric acid / calcium carbonate experiment previously described.

The results in the problem list were used to draw Graph 4, which indicates the volume of oxygen liberated from hydrogen peroxide in the presence of a little manganese oxide catalyst. Tangents have been drawn to determine the rate of reaction, initially and after 50 seconds. For ease of calculation these should be drawn by choosing the denominator (rather than the numerator), a convenient number to facilitate division, but this is less important if calculators are available.

Footnote Chemists, remember chords cut graphs twice, tangents touch once.

Chemists will want to discuss how and why in this experiment the reaction rate varies and how the graph is altered if no catalyst is used. Mathematicians may find the slant given to the data in the appendix more relevant.

Two intersecting straight lines

Sometimes the intersection of two straight lines has a special scientific significance. The difficulty in chemistry is to obtain enough results of sufficient accuracy quickly so that the pupils are convinced that two straight lines better represent the results than a curve (see Graph 5a). The appendix contains the results of two such experiments which determine the relative proportion of substances, measured in moles, in which two substances react. In the one, different amounts of sulphuric acid and sodium hydroxide solutions are mixed together and the temperature increase measured. The amounts of the two chemicals used are measured by the volumes of their solutions such that the *total* volume remains constant. The purpose of the experiment is to discover the simple whole number proportion in which sulphuric acid and sodium hydroxide react to produce the maximum temperature rise. One could structure the experiment so that maximum temperature rise occurred at one of the readings taken, but it is more realistic for it to be between two points as in Graph 5b. This shows 0.0335 moles sodium hydroxide react with 0.0165 moles sulphuric acid, i.e. 2 moles : 1 mole respectively.

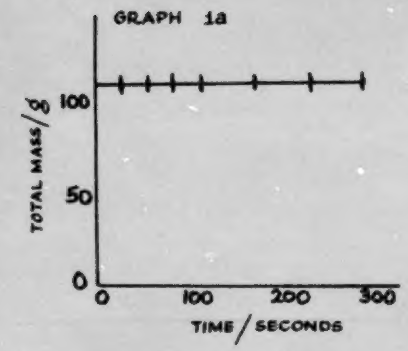
Edited by Alan Rogerson

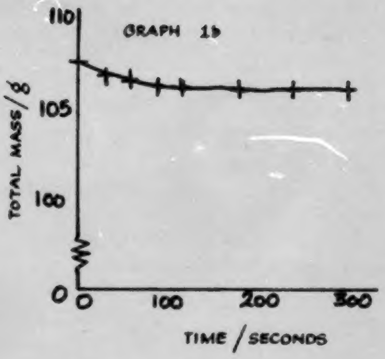
④

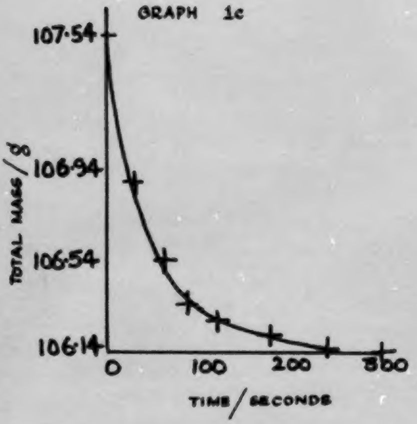
Mathematics and Chemistry: The Classroom Interface

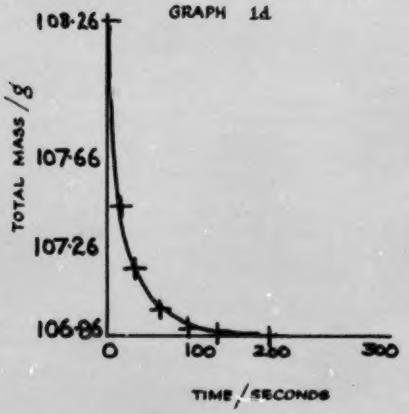
Gordon Woods

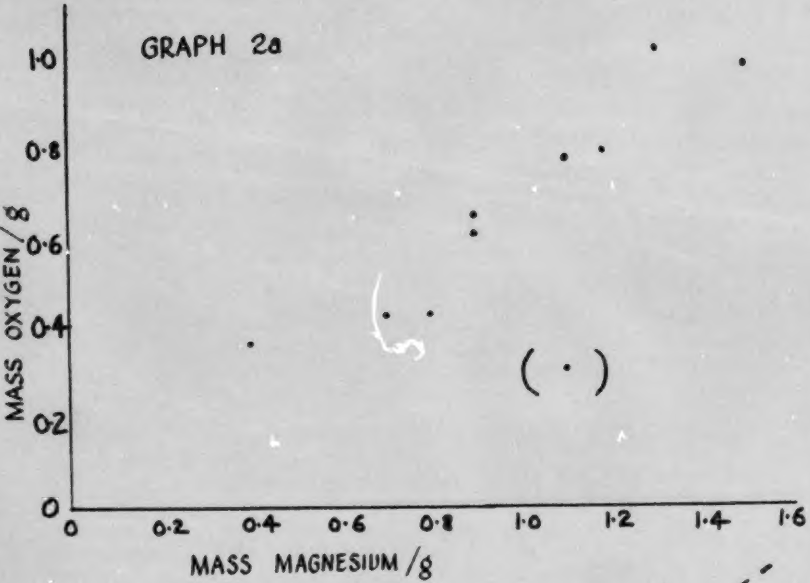
∫y dx
-b±√b²-4ac
2a
010
001
≥
÷
×
C₂H₅OH
E=mc²
NaOH

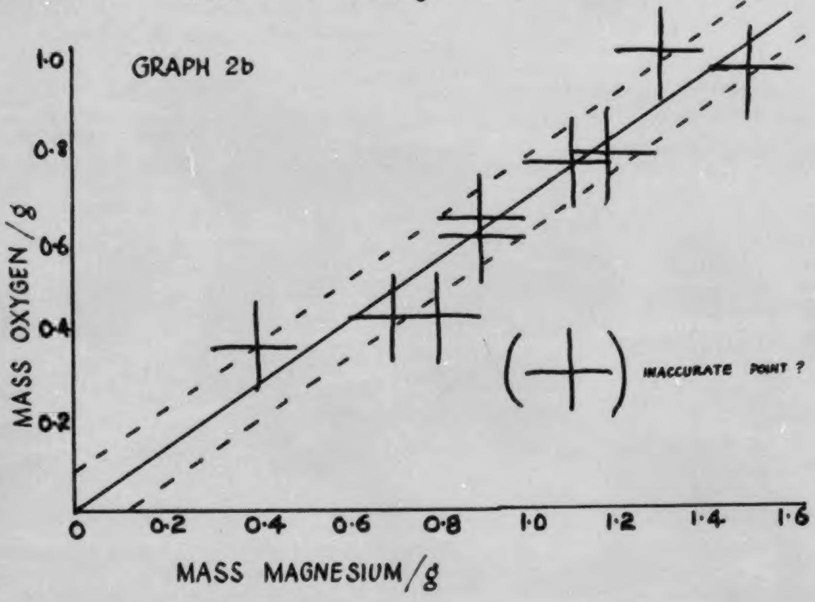

GRAPH 1a
TOTAL MASS / g
100
50
0
0
100
200
300
TIME / SECONDS

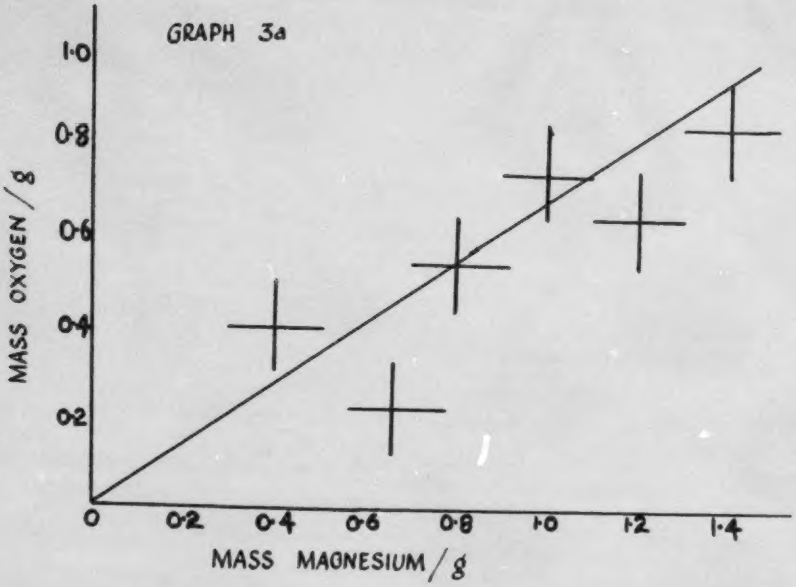

GRAPH 1b
TOTAL MASS / g
110
105
100
0
0
100
200
300
TIME / SECONDS

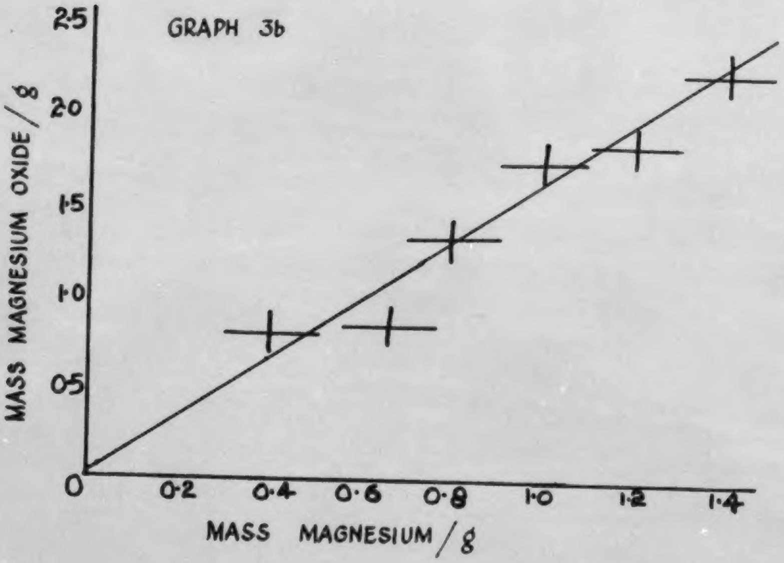

GRAPH 1c
TOTAL MASS / g
107·54
106·94
106·54
106·14
0
100
200
300
TIME / SECONDS

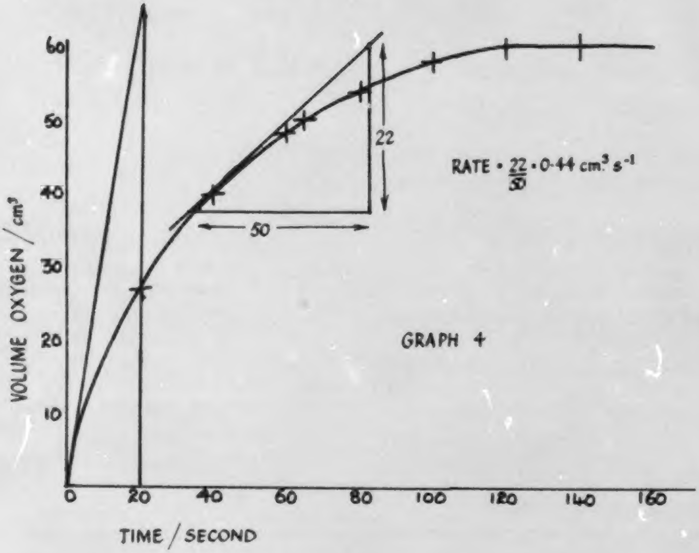

GRAPH 1d
TOTAL MASS / g
108·26
107·66
107·26
106·86
0
100
200
300
TIME / SECONDS

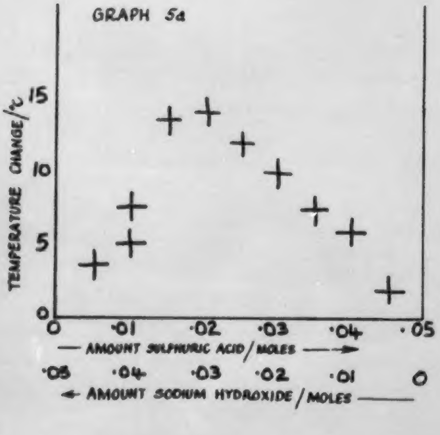

GRAPH 2a
MASS OXYGEN /g
1·0
0·8
0·6
0·4
0·2
0
0
0·2
0·4
0·6
0·8
1·0
1·2
1·4
1·6
MASS MAGNESIUM /g
(·)

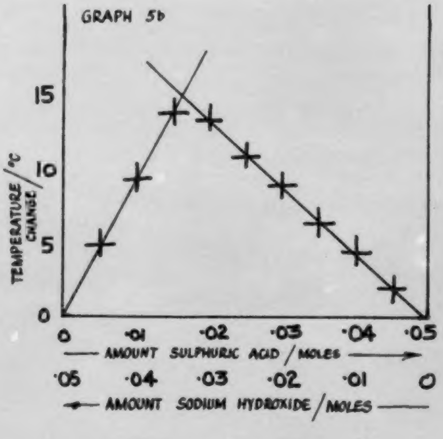

GRAPH 2b
MASS OXYGEN /g
1·0
0·8
0·6
0·4
0·2
0
0·2
0·4
0·6
0·8
1·0
1·2
1·4
1·6
MASS MAGNESIUM /g
(+) INACCURATE POINT ?


GRAPH 3a
MASS OXYGEN / g
1·0
0·8
0·6
0·4
0·2
0
0·2
0·4
0·6
0·8
1·0
1·2
1·4
MASS MAGNESIUM / g


GRAPH 3b
MASS MAGNESIUM OXIDE / g
2·5
2·0
1·5
1·0
0·5
0
0·2
0·4
0·6
0·8
1·0
1·2
1·4
MASS MAGNESIUM / g


VOLUME OXYGEN / cm³
60
50
40
30
20
10
0
22
50
RATE = $\frac{22}{50}$ = 0·44 cm³ s⁻¹
GRAPH 4
0
20
40
60
80
100
120
140
160
TIME / SECOND


GRAPH 5a
TEMPERATURE CHANGE / °C
15
10
5
0
0
·01
·02
·03
·04
·05
— AMOUNT SULPHURIC ACID / MOLES →
·05
·04
·03
·02
·01
0
← AMOUNT SODIUM HYDROXIDE / MOLES —


GRAPH 5b
TEMPERATURE CHANGE / °C
15
10
5
0
0
·01
·02
·03
·04
·05
— AMOUNT SULPHURIC ACID / MOLES →
·05
·04
·03
·02
·01
0
← AMOUNT SODIUM HYDROXIDE / MOLES —

Section 2

The function of this Section is to offer specific suggestions to the practising chemistry class teacher of those areas where new mathematics is *positively* helpful. It is hoped this may catalyse further contact between chemistry and mathematics colleagues in a particular school, for this is where, primarily, cooperation is needed. The chemistry teacher can take advantage of the language and symbolism already learnt in the mathematics class, so it is important to establish when particular mathematical techniques are taught.

First, it must be stressed that the mathematical ideas (the content of which 12-15 year old pupils will meet in the mathematics class) are at a relatively low mathematical level. Hence the pupil should be able to transfer effectively his mathematical techniques to the chemistry lesson. This means not only numeracy skills, but also understanding chemical concepts. As most of the applications involve diagrams, they are particularly helpful to the pupil who is not verbally fluent. Furthermore the low level of some of the new mathematical ideas enable them to be assimilated by both chemistry teachers (and pupils) who have been nurtured on a traditional mathematical diet!

Venn diagrams and sets

In most countries sets and Venn diagrams are met early in today's mathematics courses. The Venn diagram shows logical ideas in a visual form. A common type of Venn diagram is seen in diagrams 1 and 2. Each diagram depicts two sets A and B of objects, where each object of A is also an object of B. The text below each diagram explains the details. The example illustrated in diagram 1 is well-known to the students from their work in mathematics. The example illustrated in diagram 2 belongs to the domain of chemistry.

1.

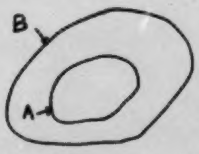


Diagram 1

A: The set of all equilateral triangles.

B: The set of all triangles.

2.

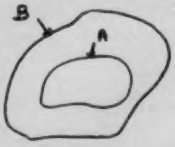


Diagram 2

A: The set of all alkalis

B: The set of all bases.

The student will from his mathematics lessons know that a set A is said to be included in a set B (or to be a subset of B) if, and only if, each element of A is also an element of B. This quite "natural" way of speaking can now easily be applied to the chemical context.

The symbol $\{\ \}$ is used to denote a 'set' of objects, usually with some common property. The symbol $\subset$ is used for 'is included within' with the open side directed towards the more general set.

$\{\text{alkenes}\} \subset \{\text{hydrocarbons}\}$

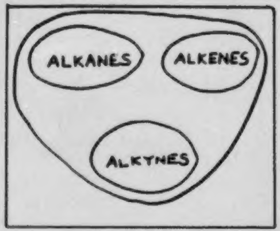
$\{\text{primary alcohols}\} \subset \{\text{alcohols}\}$

$\{\text{alkanes}\} \subset \{\text{hydrocarbons}\}$

$\{\text{electrolytes}\} \subset \{\text{conductors}\}$ etc.

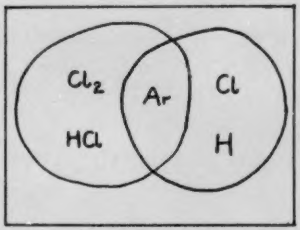
Naturally each of these statements could be represented by diagram 1, or the middle two statements could be combined as part of diagram 3. Note: there is no need to use circles to represent sets – any closed curve not cutting itself will do. The curves are often enclosed in a rectangle which represents the *universal set*, which in many chemical examples is the set of all substances.

The mathematician may have been surprised to see the word 'object' used where he might have used '*elements*' for members of a set. This was to avoid the linguistic confusion created where chemists and mathematicians use precise technical terms with entirely different meanings. Element, base, solution and group are four examples of this!



{hydrocarbons}

Diagram 3



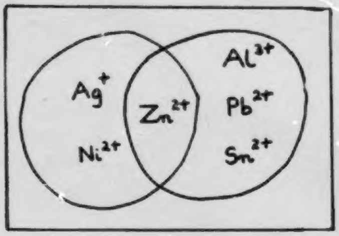
{molecules} {atoms}

Diagram 4

A second type of Venn diagram (diagram 4) represents two classifications which are partially overlapping. An atom is sometimes defined as the smallest part of an element which can take part in a chemical reaction, and a molecule as the smallest part of a substance which can exist independently. Since a single atom of the gas argon exists independently, it is also correct to call this a molecule. Hence argon (chemical symbol Ar) is common to both sets.

If the left-hand set in diagram 4 had referred to elements with metal properties and the right-hand set had referred to elements with non-metal properties then the overlap portion would be *metalloids*, elements possessing properties typical of both metals and non-metals. Similarly *amphoteric* substances can both react as acids and bases.

Other examples are non-metal oxides and acid anhydrides, ionisation and dissociation, saturated organic compounds and alkanes, metal hydroxides soluble in sodium hydroxide and ammonia solution (diagram 5).



cations forming hydroxides soluble in ammonia

cations forming hydroxides soluble in sodium hydroxide

Diagram 5

This last example is not just a summary of observation but arises from the metalloid nature of the elements whose hydroxides dissolve in excess sodium hydroxide as opposed to the ability of transition metal ions to form complexes with the weak base ammonia. The presence of Zn^{2+} in both classifications is clear justification for its position in the periodic table.

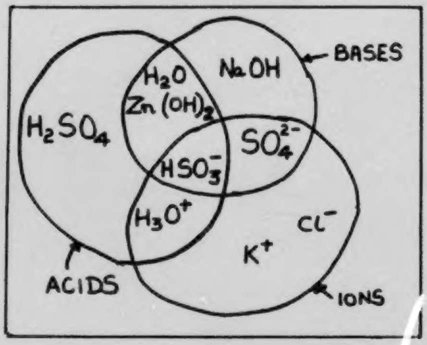


Diagram 6

Diagram 6 represents a still more complex situation of 3 partially overlapping classifications, of ions, acids, and bases, all of which apply to HSO_3^-. All these examples should stimulate the pupil to think of the similarity and differences between the classifications. He should be encouraged to put scientific examples on the diagrams.

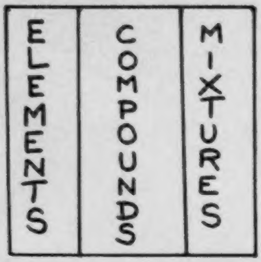


Diagram 7a

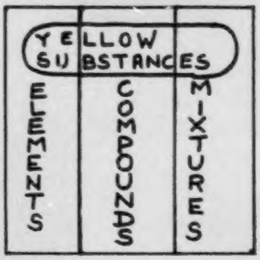


Diagram 7b

Diagram 7a, represents all substances on earth, and indicates that any substance must belong to one of 3 mutually exclusive sets. Diagram 7b, could be used as an aid to discussion, challenging pupils to put examples in the six classes.

Chemists use a diagram for the electronic structure of molecules with covalent bonds which is outwardly similar to a Venn diagram, although the mathematician will see significant differences. For example young pupils recognise the overlap portions of diagram 8 which contain a pair of electrons as being a single covalent bond – their more sophisticated elders would prefer a chemical description involving the overlap of orbitals.

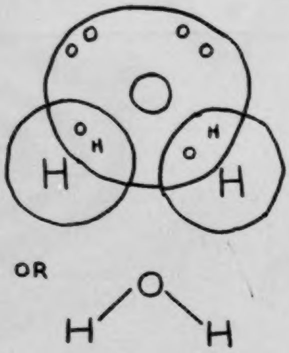


Diagram 8

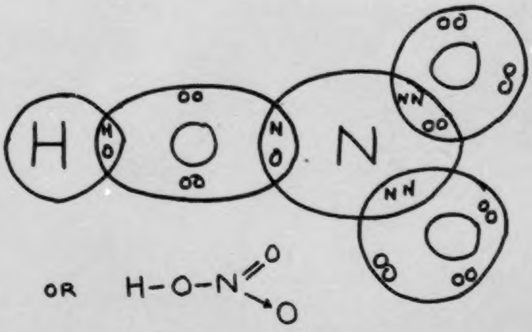


Diagram 9

The electronic structure of nitric acid (diagram 9) further illustrates the formation of a double bond (two electrons pairs in the overlap) and a dative covalent bond in which one atom, here the oxygen, provides both electrons of the shared pair. In both these diagrams the frequently used dot and cross notation for electrons has been replaced by the use of small capital letters for the elements from which the electrons originated. One wonders why this self evident notation is not more extensively used by chemists.

Thus at a very simple mathematical level, Venn diagrams provide the chemistry teacher with a means of understanding overlapping concepts or merely as a visual method of remembering factual detail. A pertinent comment on the usefulness of Venn diagrams is that at least three authors over a ten year period have *independently* discovered their usefulness in chemistry teaching and published articles about them.

Seeking a relationship

Elementary chemistry has no situations where mathematical dependence is other than direct or inverse proportion. Older school pupils however use gas syringes to study the time (t) taken for a given volume of gas, of relative molecular mass (m), to effuse from a small hole. The times are all accurate to ± 1 second.

Time (t) / seconds	10	15	29	39	48	60
Gas	H_2	He	NH_3	O_2	CO_2	Cl_2
R.M.M. (m)	2	4	17	32	44	71

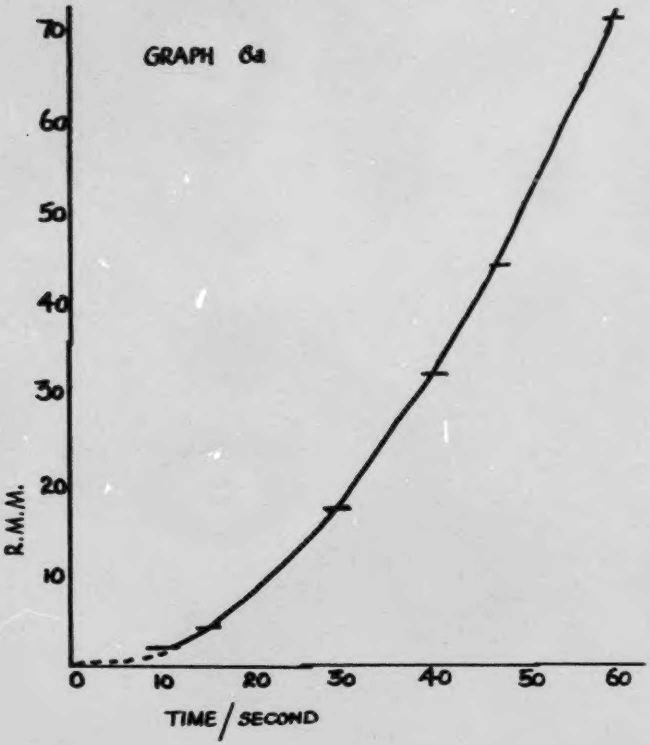
Direct inspection of the results suggests that (m) is not proportional to (t) as confirmed by Graph 6a, which is not linear, with the curve possibly going through the origin, the inaccuracy in the time measurement being proportionally more significant in that area. When pupils are challenged to suggest the kind of relationship satisfied by the data, some will correctly suggest m against t^2, or the equivalent $(m)^{\frac{1}{2}}$ against t, as plotted in Graph 6b.

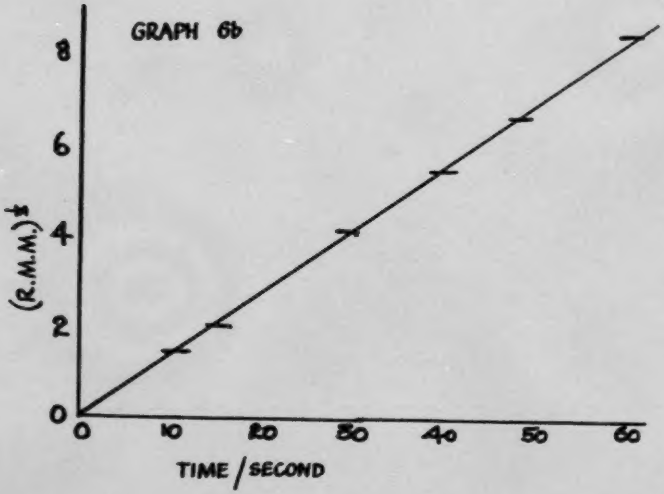
Footnote: If $t = k(m)^n$, $\log t = \log k + n \log m$ so some pupils may suggest the more general approach of plotting log m against log t, the slope being $^1/_n$. This relationship bears the name of its discoverer, Graham.

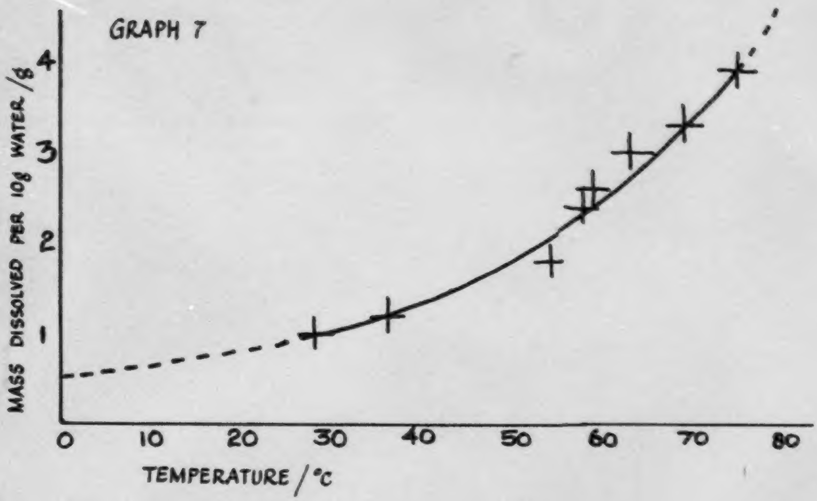
Curves not obeying any simple mathematical relationship. Graph 7 represents a selection of class results to investigate how the solubility (the mass dissolved in 100g of water) of a substance varied with temperature. Chemically this is conveniently administered using only 10 g water and by pairing off pupils, each pair being given a high mass and a low mass, and noting the temperature at which a solution becomes saturated on cooling. Do these points form a straight line or a curve? Should it go through the origin?

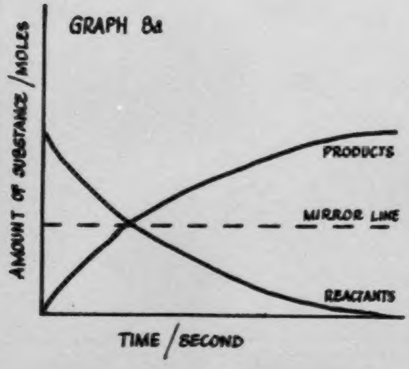
Reflections and Stretches

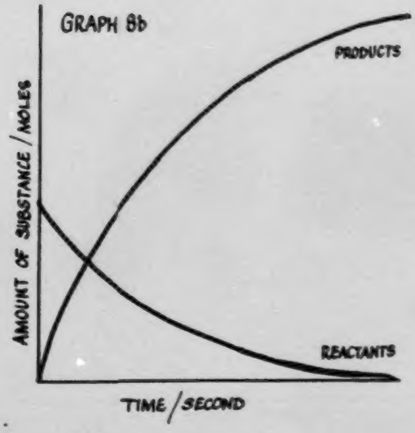
In chemical kinetics the amounts of reacting substances (reactants) and products are studied as a function of time. In a reaction for which there are an equal number of reactant and product molecules, the amount of each substance present is found by reflection in the mid-amount line, as shown in Graph 8a. Graph 8b represents the situation when twice as many product molecules are found as reactant molecules removed, hence the reflection is followed by a ×2 one way stretch parallel to the amount axis.


GRAPH 6a
R.M.M.
70
60
50
40
30
20
10
0
10
20
30
40
50
60
TIME / SECOND


GRAPH 6b
(R.M.M.)½
8
6
4
2
0
0
10
20
30
40
50
60
TIME / SECOND


GRAPH 7
MASS DISSOLVED PER 10g WATER /g
1
2
3
4
0
10
20
30
40
50
60
70
80
TEMPERATURE / °C


GRAPH 8a
AMOUNT OF SUBSTANCE / MOLES
PRODUCTS
MIRROR LINE
REACTANTS
TIME / SECOND


GRAPH 8b
AMOUNT OF SUBSTANCE / MOLES
PRODUCTS
REACTANTS
TIME / SECOND

Arrow (or mapping) diagrams

Students in the mathematical classroom draw arrow diagrams (diagram 10a), which show the relation between mathematical elements in two sets. The arrow in this particular diagram represents 'is a factor of', hence both 2 and 4 are factors of 8. The chemistry student could draw a similar diagram (diagram 10b) which, without using or explaining specific mathematical terms, clearly shows that aqueous solutions containing two of the ions listed are green, and that none of the listed ions have orange coloured solutions.

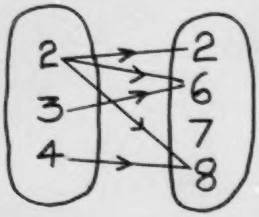


Diagram 10a

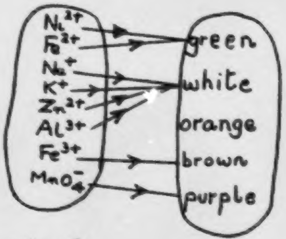

Ni^{2+}
Fe^{2+}
Na^{+}
K^{+}
Zn^{2+}
Al^{3+}
Fe^{3+}
MnO_4^-
green
white
orange
brown
purple
formula of ion
colour

Diagram 10b

Diagram 11a depicts the chemical concept of isomerism, in which the same molecular formula applies to two or more substances with different structural formulae. The mathematician would describe this as a *many to one* mapping, which is (in contrast to diagram 11b), a *one to one* mapping. Diagram 11b shows the relation between the same 6 substances and their names. Curiously enough the mathematician would call the 6 substances elements (of a set) but the chemist would describe them as compounds and NOT elements!

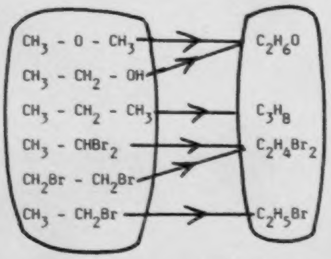

$CH_3 - O - CH_3$
$CH_3 - CH_2 - OH$
$CH_3 - CH_2 - CH_3$
$CH_3 - CHBr_2$
$CH_2Br - CH_2Br$
$CH_3 - CH_2Br$
C_2H_6O
C_3H_8
$C_2H_4Br_2$
C_2H_5Br
structural formula
molecular formula

Diagram 11a

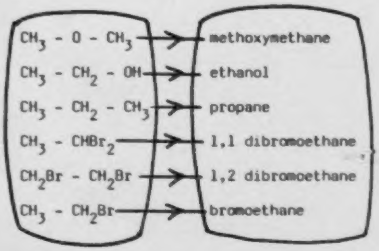

$CH_3 - O - CH_3$
$CH_3 - CH_2 - OH$
$CH_3 - CH_2 - CH_3$
$CH_3 - CHBr_2$
$CH_2Br - CH_2Br$
$CH_3 - CH_2Br$
methoxymethane
ethanol
propane
1,1 dibromoethane
1,2 dibromoethane
bromoethane
structural formula
name

Diagram 11b

Such arrow diagrams can equally be used to relate substances to their physical or chemical properties, to stress whether the property is unique to that substance or more general.

Flow diagrams

Flow diagrams are used in mathematics for many purposes including an introduction to the language of computers. They diagrammatically depict a series of systematically related actions, usually shown in rectangular *operation boxes*. Diamond shaped *decision boxes* are used where a question needs to be answered before selecting one of two routes, depending on whether Yes or No is the answer to the question. *Start and stop boxes* are conventionally drawn as ellipses.

Flow diagrams can be constructed for such sequences as dialing a telephone number or assembling a motor car but their chemical uses are mainly in three fields

(1) Analysis
(2) Steps in practical work
(3) Classification procedure.

The two examples chosen to illustrate the use of flow diagrams in analysis are an elementary approach to the identity of a cation by addition of sodium hydroxide solution (diagram 12) and, at a more advanced level, the identification of an anion, which gives a precipitate with silver nitrate solution (diagram 13). Diagram 12 is simplified to suit the needs of a chemist; in strict mathematical terms the identification of a particular ion from its colour results from a series of questions to which the answer is Yes or No, e.g. Is the precipitate green? Is the precipitate brown? etc.

Diagrams 14 and 15 illustrate the use of flow diagrams in setting out the sequence of steps in practical work. Diagram 14 shows the formation of a soluble salt by neutralisation using an indicator prior to meeting a titration technique, and diagram 15 describes the more advanced experiment of estimation of the oxidizing agent, potassium iodate (KIO_3), by titration against sodium thiosulphate solution in the presence of excess potassium iodide solution. Pupils can take traditional (cookery book) instructions for an experiment and devise a flow diagram to obtain an overall view of the experiment through understanding the necessary sequence of operations.

After being taught how experiments lead to the classification of substances into mixtures, compounds and elements, either metal or non-metal, more able pupils will be able to link the procedures underlying this classification in a flow diagram (like diagram 16). For most other pupils the flow diagram will need to be *given* them to summarise and link work which they regarded as separate units. A similar experimental procedure involving the attempted passage of an electric current through a substance, first as a solid, then when molten enables one to classify a solid by its lattice type.

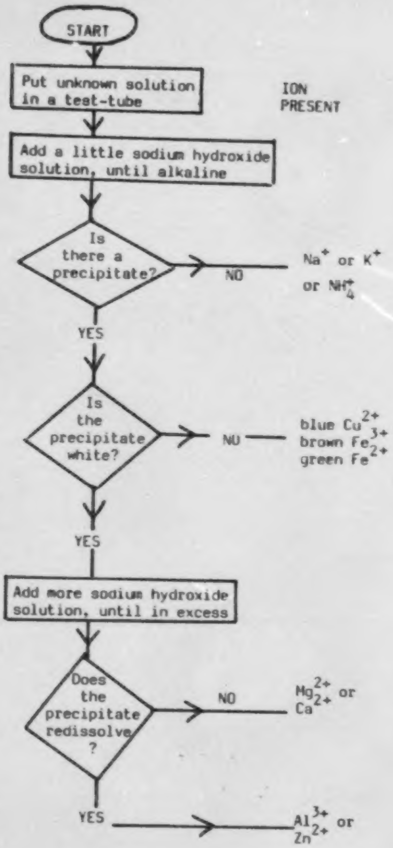

START
Put unknown solution in a test-tube
ION PRESENT
Add a little sodium hydroxide solution, until alkaline
Is there a precipitate?
NO
Na^+ or K^+ or NH_4^+
YES
Is the precipitate white?
NO
blue Cu^{2+}
brown Fe^{3+}
green Fe^{2+}
YES
Add more sodium hydroxide solution, until in excess
Does the precipitate redissolve ?
NO
Mg^{2+} or Ca^{2+}
YES
Al^{3+} or Zn^{2+}

Diagram 12

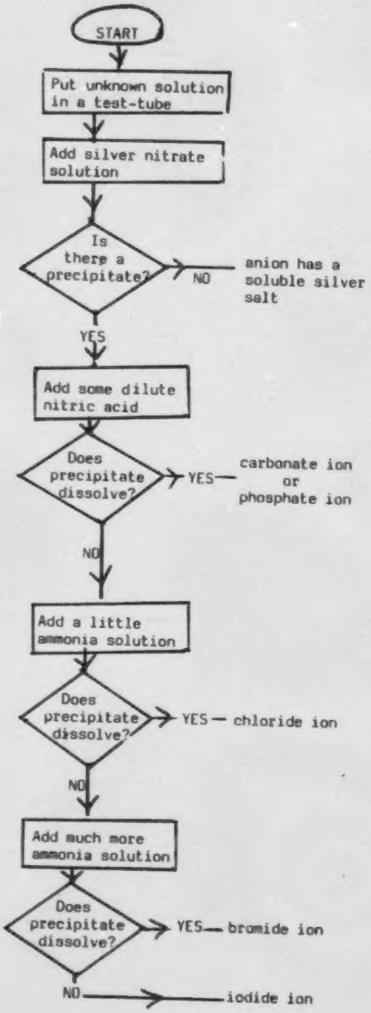

START
Put unknown solution in a test-tube
Add silver nitrate solution
Is there a precipitate?
NO
anion has a soluble silver salt
YES
Add some dilute nitric acid
Does precipitate dissolve?
YES
carbonate ion or phosphate ion
NO
Add a little ammonia solution
Does precipitate dissolve?
YES
chloride ion
NO
Add much more ammonia solution
Does precipitate dissolve?
YES
bromide ion
NO
iodide ion

Diagram 13

Performing an iodine-thiosulphate titration

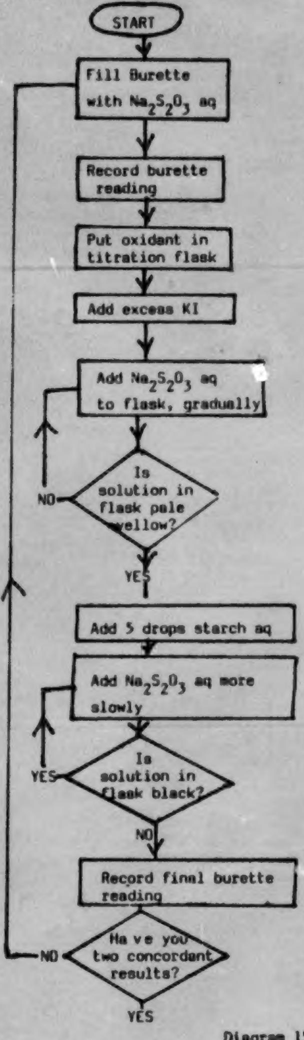

START
Fill Burette with $Na_2S_2O_3$ aq
Record burette reading
Put oxidant in titration flask
Add excess KI
Add $Na_2S_2O_3$ aq to flask, gradually
Is solution in flask pale yellow?
NO
YES
Add 5 drops starch aq
Add $Na_2S_2O_3$ aq more slowly
Is solution in flask black?
YES
NO
Record final burette reading
Have you two concordant results?
NO
YES

Diagram 15

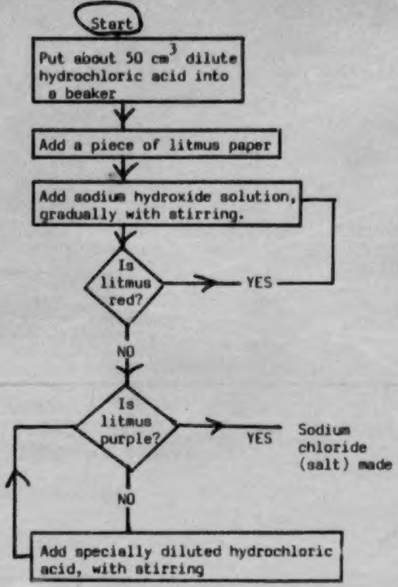

Start
Put about 50 cm^3 dilute hydrochloric acid into a beaker
Add a piece of litmus paper
Add sodium hydroxide solution, gradually with stirring.
Is litmus red?
YES
NO
Is litmus purple?
YES
Sodium chloride (salt) made
NO
Add specially diluted hydrochloric acid, with stirring

Diagram 14

Classifying a substance

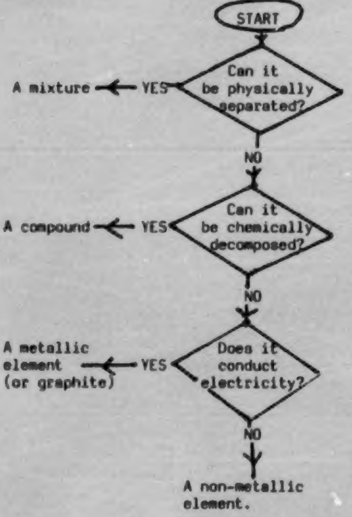

START
Can it be physically separated?
YES
A mixture
NO
Can it be chemically decomposed?
YES
A compound
NO
Does it conduct electricity?
YES
A metallic element (or graphite)
NO
A non-metallic element.

Diagram 16

Classifying the lattice type of a solid.

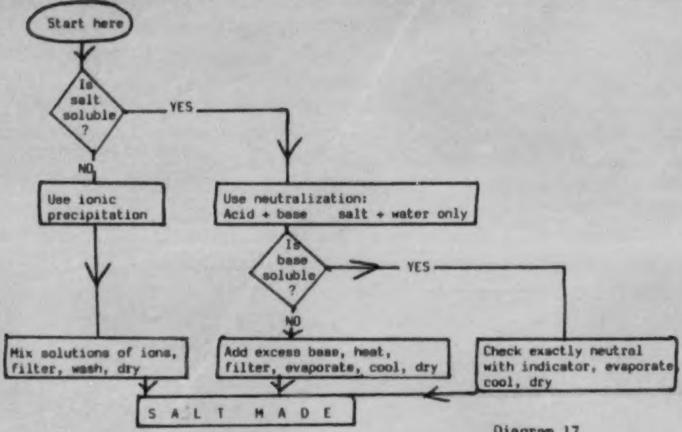

Start here
Is salt soluble ?
YES
NO
Use ionic precipitation
Use neutralization: Acid + base salt + water only
Is base soluble ?
YES
NO
Mix solutions of ions, filter, wash, dry
Add excess base, heat, filter, evaporate, cool, dry
Check exactly neutral with indicator, evaporate, cool, dry
SALT MADE

Diagram 17

Diagram 17 briefly describes three ways of preparing a salt, one and only one of which is suitable for a particular salt according to the answers to two questions about the solubility of the salt itself, and the base from which it can be derived. These questions must be asked in the correct order. Their answers may be found from published data.

N.B. At a more advanced level one must add a fourth method for those salts which react with water, direct combination from the elements being used in most cases. How would the diagram be altered to accommodate this extra method?

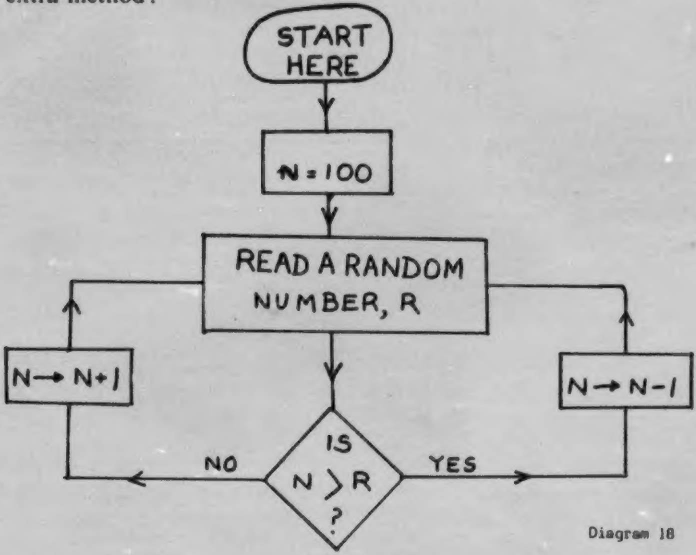

START HERE
N = 100
READ A RANDOM NUMBER, R
N → N+1
N → N-1
IS N > R ?
NO
YES

Diagram 18

The phenomenon of diffusion can be illustrated or introduced using the programme shown in diagram 18. Consider two vessels of equal volume joined by a narrow tube and closed with a tap. One vessel, A, initially contains 100 molecules and the other, B, is empty. The tap is now opened. The question in the decision box is a measure of the probability that the next molecule going from one box to the other will go from A to B. The individual random numbers are continuously generated by reading the final two numbers from a list of numbers from a telephone directory. Thus initially when $n = 100$, $n > r$, so that the molecule is certain to move from A to B. Naturally the probability gradually decreases but the programme has no specific end. The pupil stops when he realises that he is getting random variations about $n = 50$.

A graph is then plotted of m against the total number of random numbers generated. Larger numbers of molecules, say $m = 500$, can be studied as a group exercise. How could this mathematical model, for such it is, be extended to consider different sized variables? Because pupils regard the programme as a 'fun' exercise, they can first carry it out and then be introduced to the theory later according to their mathematical development.

Topology and Isomerism

Topology is concerned with the relative arrangements of points in a geometrical figure. The points where *arcs* (straight or curved lines) meet are called *nodes*. The *node order* refers to the number of arcs at a point. Two figures are topologically equivalent if one can be changed into the other by stretching and twisting only, as if the figures were drawn on a very stretchy piece of rubber, but without cutting or joining.

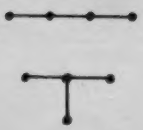


Diagram 19

The two structural isomers of formula C_4H_{10} are often a pupil's introduction to isomerism. To help him, the teacher can draw the carbon skeletons in an overtly mathematical manner (diagram 19), stressing the role of the carbon atoms as nodes by large dots. It is now immediately apparent that these representations are topologically different although they have the same number of nodes and arcs, four and three in this example, the criterion for such isomerism.

In a similar manner diagram 20 (a-c) represents the three isomers of formula C_5H_{12}. Formulae (d-f) are commonly produced by pupils as 'extra isomers', but they should be rejected because they are topologically equivalent to c. Most chemistry teachers probably encourage their pupils to write the longest carbon chain across the page but nevertheless the teacher is often shown other 'extra isomers' like (g-i) which must again be rejected because they are topologically equivalent to one of (a-c).

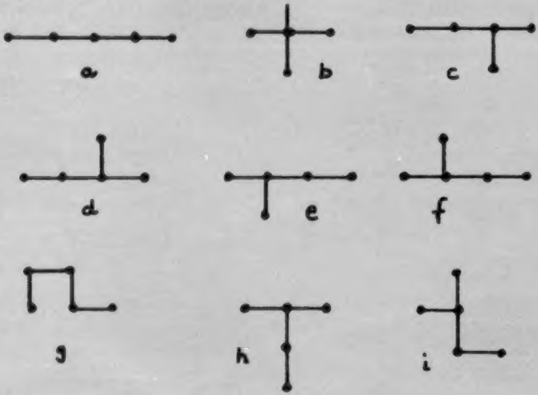


Diagram 20

The isomers of formula C_6H_{14} (diagram 21 (a-e)) underline that it is topological equivalence which is the criterion for two molecules being the same. Formulae (f and g) are common alternative representations of (e), to which they are topologically equivalent.

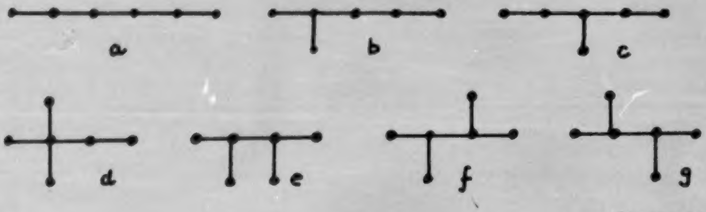


Diagram 21

These ideas have been extended to larger alkane molecules to permit calculation of the number of structural isomers possible. Incorrect answers were obtained until iterative techniques were developed in the early 1930s.

Molecular models are a most valuable aid to the pupil's understanding of the concepts of isomerism, particularly if they handle the models. These mathematical ideas concerning carbon skeleton isomerism provide an alternative or additional approach, with the advantage that they can be written down either on the blackboard, by the teacher, or in the examination room, by the pupil, where molecular models are likely to be excluded.

Simple Probability

Students of new mathematics are introduced to probability at an early stage so they would easily cope with obtaining a mass spectrum for chlorine molecules (Cl_2) given that the individual chlorine atoms are of two different mass numbers, 35 and 37, in relative abundance, $\frac{3}{4}$ and $\frac{1}{4}$ respectively. Thus

abundance $^{35}Cl - ^{35}Cl$ molecule $= \frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$
,, $^{35}Cl - ^{37}Cl$,, $= 2 \times \frac{3}{4} \times \frac{1}{4} = \frac{6}{16}$
,, $^{37}Cl - ^{37}Cl$,, $= \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$

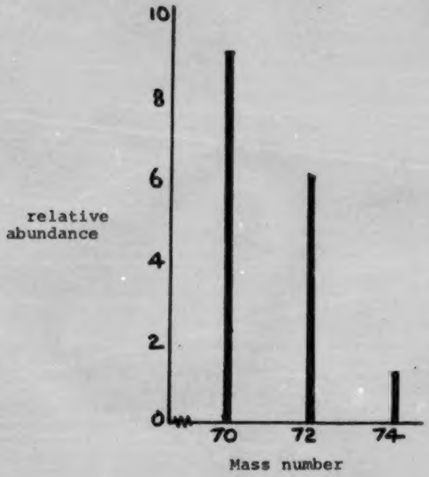


Diagram 22

Diagram 22 shows the mass spectrum obtained.

More advanced pupils will be able to determine more complex formulae from mass spectra, as is done in chemical research.

Postscript

This booklet has presented various ways in which 'modern mathematics' can be positively helpful to the chemistry teacher – first in areas of traditional skills and then in new ideas and notation. It is hoped that this will lead in turn to an increased dialogue between mathematics and chemistry teachers, and hence to more cooperation at school level.

Problem List

This selection of 24 questions is arranged according to the mathematical content and indicates the skills a chemistry teacher requires of his pupils. Apart from the few (specifically mentioned and chosen to illustrate particular mathematical techniques) the questions cover chemistry met up to the age of 16 in Great Britain.

Accuracy

1. A chemical balance records the mass of a substance to the nearest one hundredth of a gram. (*a*) If the balance reads 2.22 g, within what range does the mass lie? (*b*) What mass should be recorded if the balance records, to its particular accuracy, (i) just 2 g, (ii) 2.2 g?

Four Rules and Brackets

2. Calculate the relative molecular masses of the following substances by adding up their relative atomic masses. Note the subscripts after the brackets multiply everything inside the bracket.
   (i) CO_2, (ii) $NaHCO_3$, (iii) $Ca(OH)_2$, (iv) $Ca(HCO_3)_2$, (v) $Na_2CO_3.10(H_2O)$.
   The relative atomic masses for single atoms are $H = 1$, $O = 16$, $Na = 23$, $Ca = 40$.

Ratio and Proportion

3. (*a*) One mole of zinc atoms has a mass of 65 g. How many moles of zinc are there in (i) 13 g zinc, (ii) 72 g zinc, (iii) 0.46 g zinc?

   (*b*) One mole of copper atoms has a mass of 64 g. How many moles of copper atoms are there in (i) 16 g copper, (ii) 400 g copper, (iii) 0.6 g copper?

4. (*a*) $C + O_2 \rightarrow CO_2$. This chemical equation shows that 12 g carbon (C) produce 44 g carbon dioxide (CO_2). What mass of carbon is needed to make (i) 8.8 g CO_2, (ii) 0.22 g CO_2, (iii) 1 kg CO_2?

   (*b*) $SO_3 + H_2O \rightarrow H_2SO_4$. This chemical equation shows that 80 g SO_3 are needed to produce 98 g H_2SO_4. What mass of SO_3 is needed to produce (i) 49 g H_2SO_4, (ii) 2.45 g H_2SO_4, (iii) 180 g H_2SO_4, (iv) 0.2 g H_2SO_4?

5. The following results were obtained in an experiment to determine the formula of a sample of copper oxide.

Mass porcelain boat empty = 16.96 g
,, ,, ,, + copper oxide = 19.03 g
,, ,, ,, + copper = 18.80 g

Calculate (i) the mass of copper,
(ii) the mass of oxygen in the copper oxide,
(iii) the number of moles of copper atoms in (i), if one mole of copper atoms has a mass of 64 g,
(iv) the number of moles of oxygen atoms in (ii), if one mole of oxygen atoms has a mass of 16 g,
(v) the ratio of moles of copper atoms : moles of oxygen atoms, using small integers.

6. 64 g SO_2 contains 6×10^{23} molecules.

(*a*) How many SO_2 molecules are there in (i) 6.4 g SO_2, (ii) 0.064 g SO_2, (iii) 128 g SO_2, (iv) 3.2 g SO_2, (v) 1.6 mg SO_2, i.e. 1.6×10^{-3} g SO_2?

(b) What is the mass of SO_2 containing (i) 12×10^{23} molecules, (ii) 1.5×10^{22} molecules, (iii) 10^{20} molecules, (iv) one molecule?

7. Two compounds between iron and oxygen only contain respectively 70.0% and 77.7% by mass of iron. Show that the masses of iron which combine with a fixed mass of oxygen are in simple whole number proportion. (What are convenient choices for the fixed mass?)

Percentages

8. The results obtained in an experiment, in which some magnesium was heated in an open crucible, were:

Mass of crucible empty = 9.88 g
,, ,, ,, + magnesium ribbon = 10.45 g
,, ,, ,, + 'magnesium' after heating = 10.83 g

Calculate (i) the mass of magnesium ribbon,
(ii) the increase in mass on heating magnesium,
(iii) the increase in mass as a percentage of the original mass of magnesium.

9. Calculate the percentage by mass of the mentioned element in the following substances. Remember that the subscript after the brackets multiplies everything within the brackets, as in Question 2.
(i) H in H_2O, (ii) K in $KHCO_3$, (iii) N in NH_4NO_3 (iv) O in $Ca(OH)_2$, (v) O in $Ca(NO_3)_2$. The relative atomic masses are H = 1, C = 12, O = 16, N = 14, K = 39, Ca = 40.

Weighted Average

10. Calculate the relative atomic mass of the element Gallium if 60% of its atoms are of mass 69 and the rest of mass 71.

Formulae

11. Rearrange the equation below to make V_2 the subject.

(*a*) $\frac{P_1V_1}{T_1} = \frac{P_2V_2}{T_2}$

(*b*) Use the above equation to obtain the value of V_2 when
$V_1 = 45\,cm^3$, $P_1 = 105\,kN/m^2$, $T_1 = 290\,K$, $T_2 = 273\,K$, $P_2 = 101\,kN/m^2$.

Quadratic Equations (more advanced school chemistry)

12. The number of moles of HI present (2x) at equilibrium for the reaction represented by the chemical equation $H_2 + I_2 \rightarrow 2HI$ is obtained under different conditions by solving the equations that follow. In each case solve the equation for x and explain which of the two values of x is physically possible. (The bracketed terms in the denominator represent the equilibrium amounts of H_2 and I_2).

(*a*) $64 = \frac{(2x)^2}{(1 - x)^2}$ (*b*) $40 = \frac{(2x)^2}{(1 - x)(2 - x)}$

13. The hydrogen ion concentration, h, in a particular ethanoic acid solution is found by solving the equation

$$2 \times 10^{-5} = \frac{h^2}{(1 - h)\,10}$$

Hint: Is h large or small? What reasonable approximation can be made to simplify the calculation?

Graphs

A simple linear graph

14. The table of values lists the mass of oxygen that was found to react with different masses of zinc to form zinc oxide. Plot the mass of oxygen (y axis) against the mass of zinc (x axis).

Mass oxygen / g	0.12	0.20	0.32	0.40	0.48
Mass zinc / g	0.49	0.81	1.30	1.63	1.95

(*a*) Within experimental error, do the points lie on a straight line? What mass of oxygen reacts with exactly 1 g of zinc?

(*b*) An important chemical law states that any compound always contains the same elements in the same proportion by mass. Explain how these results illustrate this.

Choice of scale and axis

15. When calcium carbonate and hydrochloric acid react together they produce a gas (carbon dioxide) which is given off, hence the mass of the reaction vessel and contents decreases. The decrease, the mass of carbon dioxide, is the same for both sets of experimental data below and is small compared to the total masses recorded. How can the scale and range of values best be chosen to represent the data? Plot the time on the x axis so that in each case the rate of reaction can be obtained after 30 seconds

Experiment I

Mass of reaction flask and contents / g	107.54	106.88	106.56	106.36	106.28	106.19	106.16	106.14
Time / seconds	0	30	60	90	120	180	240	300

Experiment II

Mass of reaction flask and contents / g	108.26	107.40	107.16	106.98	106.90	106.87	106.86
Time / seconds	0	15	30	60	90	120	180

The experiments are the same except that they are carried out in different, but similar vessels (hence the slight mass differences at the start and end), and in the second case the calcium carbonate is in smaller sized lumps. What effect does the lump size have on (*a*) the mass of gas given off; (*b*) the rate at which gas is given off?

More advanced school chemistry and a mathematically demanding graph

16. The activation energy (E) of a chemical reaction is obtained from the slope of the linear graph obtained when log k is plotted against 1 / T, where k is the rate constant and T the temperature. Find E, which equals the gradient × 4.6. Be careful in your choice of scale for the 1 / T axis as the values are over a small range. An electronic calculator is very helpful to provide a table of values of log k and 1 / T from the listed experimental values.

k / $mol^2 dm^6 s^{-1}$	4.5×10^{-5}	2.5×10^{-3}	6.4×10^{-2}	1.35
T / K	556	629	700	780

More advanced school chemistry involving logs

17. The rate constant for the isomerisation of cyclopropane is the gradient × 2.3 when the log (percentage of cyclopropane left) is plotted against time.

Find the rate constant from the data below.

Time / hours	0	2	5	10	20	30
%age cyclopropane left	100	91	79	63	40	25

Mathematical note. The 2.3 is ln 10 ($\log_e 10$)

Intersection of two linear graphs

18. Find the proportions, using whole numbers less than five, in which moles of barium hydroxide solution and moles of iron (III) chloride solution react. This is determined from the *minimum* number of moles of iron (III) chloride which are required to produce a *maximum* amount of precipitate with a fixed number of moles of barium hydroxide (1×10^{-3}).

    Plot the amount of precipitate (y axis), expressed as its height, against the number of moles of iron (III) chloride (x axis)

Amount of precipitate / cm	0	0.55	1.07	1.60	1.81	1.81	1.81
Moles of iron (III) chloride, all $\times 10^{-4}$	0	2	4	6	8	10	12

19. Determine the proportion, expressed in small whole numbers of moles, in which sulphuric acid and sodium hydroxide react using the experimental information below. The correct proportion produces the greatest temperature rise, but this is not at any of the points measured.

    Plot the temperature rise (y axis) against the number of moles on the x axis, the moles of sulphuric acid increasing from left to right and the moles of sodium hydroxide increasing from right to left. Do not forget to consider the temperature rise which would be obtained if one of the chemicals had been used.

Moles of sulphuric acid	0.05	0.10	0.15	0.20	0.25	0.30	0.40
Moles of sodium hydroxide	0.45	0.40	0.35	0.30	0.25	0.20	0.10
Temperature rise / °C	5.0	9.5	14.0	13.5	11.0	9.0	4.5

Rates of reaction (see also question 15)

20. When hydrogen peroxide is decomposed it produces oxygen gas. The data below lists the volume of oxygen evolved after various times. Plot the volume against the time (x axis).

    (*a*) Draw tangents to determine the rate of production of oxygen after (i) 0 seconds, (ii) 40 seconds, (iii) 80 seconds, (iv) 120 seconds, (v) 160 seconds.

Volume oxygen / cm^3	0	27	40	48	54	58	59	60	60
Time / seconds	0	20	40	60	80	100	120	140	160

    (*b*) Sketch the graph obtained if the same total amount of oxygen were obtained but at a greater initial rate.

    (*c*) After how long is half the total volume of oxygen evolved?

21. Sodium thiosulphate solution in various concentrations was reacted with a fixed concentration of hydrochloric acid at 15°C to form a definite turbidity due to sulphur. The results obtained were:

Concentration of sodium thiosulphate	0.17	0.35	0.56	0.70
Time to reach a definite turbidity / seconds	200	100	62.5	50

Plot these results so as to give a graph of concentration of sodium thiosulphate against *rate of reaction*. What does the graph suggest?

A smooth curve

22. Draw a solubility curve, a graph of solubility (y axis) against temperature for a substance from the following data.

Solubility / g per 100 g water	6.0	13.0	22.7	37.5	60.3	75.7
Temperature / °C	0	20	40	60	80	90

(*a*) What is the solubility at 50°C?

(*b*) At what temperature is the solubility 50 g per 100 g water?

(*c*) By comparing the solubilities at the two temperatures, determine the mass of crystals obtained in cooling down from 86°C to 10°C, 100 g water which is saturated at the higher temperature.

Finding a mathematical relationship (more advanced chemical content)

23. The data below lists the time (t) taken for the same volume of different gases to escape through a small hole. The times given are accurate to ± 1 seconds but the relative molecular masses (m) are exact.

(*a*) Plot m against t. Explain why the data implies that m is *not* directly proportional to t.

(*b*) Look at the graph in (*a*) and the data and suggest how m and t are mathematically related. Now plot a suitable graph to test your suggestion. Is this linear? What conclusions can you draw?

t / seconds	10	15	29	39	48	60
R.M.M.(m):	2	4	17	32	44	71

A curve, within experimental error?

24. A small class of chemistry pupils obtained the following results for the solubility of potassium chlorate at different temperatures. It is reasonable to assume that the masses are accurate to ±0.2 g and the temperatures to ±2°C. Plot the points as crosses with arms of that length to indicate the range within which the points should lay.

Solubility / g per 10 g water	1.0	1.2	1.5	1.8	2.4	2.6	3.0	3.3	3.9
Temperature / °C	28	36	68	54	58	59	63	69	75

(*a*) Which reading appears to be most inaccurate?

(*b*) Draw a best fit line and hence read off the solubility at 30°C and 60°C.

(*c*) Explain whether it is reasonable to say that the solubility is 0 at 0°C.

Edited by Alan Rogerson

⑤

Mathematical Modelling

D. N. Burghes
and
M. S. Borrie

∫ydx
$\frac{-b\pm\sqrt{b^2-4ac}}{2a}$
0:0
001
≥
×
÷
C_2H_5OH
$E=mc^2$
NaOH

Although this booklet is primarily aimed at the mathematics teacher, it should also be of interest to all science teachers. It sets out a number of lively and varied case studies suitable for modelling treatment: retail price index, weightlifting, train timetables, population changes, shopping centres and stock control. Each situation is explored in detail and each illustrates the basic modelling flow chart developed and set out at the beginning of the booklet. Modelling is often a new way of looking at the *applications* of mathematics in the outside world and in the sciences. We hope it will, therefore, provide common ground for cooperation between mathematicians and scientists.

Contents

1 Introduction to Mathematical Modelling 3
2 Forecasting: The Retail Price Index 8
3 Handicapping: Weightlifting 14
4 Planning: Train Time Tables 17
5 Predicting: Population Changes.. 23
6 Estimating: Attractiveness of Competing Shopping Centres.. .. 31
7 Decision Making: Stocking Magazines.. 37
8 Modelling in the Classroom 41
9 References, Extra Material and Data for Related Problems.. .. 43

Mathematical Modelling

D. N. Burghes
Cranfield Centre for Teacher Services,
Cranfield Institute of Technology

&

M. S. Borrie
Department of Mathematics,
Pilgrim School,
Bedford.

1 – Introduction to Mathematical Modelling

This booklet is concerned with the use of mathematics in the formulation and solution of problems arising in the real world. The process of translating such problems from their initial context to a mathematical description, solving them, and then interpreting the solutions back in the real world situation is termed "mathematical modelling". The main stages in the process are summarised in the figure 1 below.

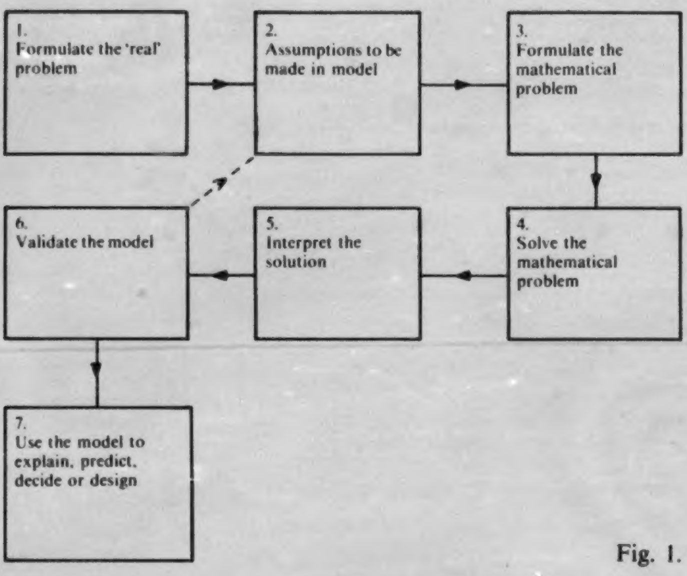

1. Formulate the 'real' problem
2. Assumptions to be made in model
3. Formulate the mathematical problem
4. Solve the mathematical problem
5. Interpret the solution
6. Validate the model
7. Use the model to explain, predict, decide or design

Fig. 1.

The left hand column represents the real world, the right hand column the mathematical world and the middle column the connection between these two worlds where firstly the problem is simplified and formalised and secondly where the mathematical solution is translated back into the real world.

In a straightforward modelling process we move from box 1 to box 7 in sequence; but most modelling is not straightforward. In many cases clues to how to solve the problem are provided by looking at the practical meaning of the mathematical expressions. We often need to spend significantly more time on particular stages; the model is sometimes not adequate for its purpose, and we may have to move from box 6 to box 2 and repeat the process using a more sophisticated model. In many cases, particularly in the social sciences, it is difficult to validate the model at all and we move straight from box 5 to box 7.

The main stages of modelling can be divided into: formulation, solution and application. The formulation stage is covered by boxes 1, 2 and 3, the solution stage by box 4 and the application stage by boxes 5, 6 and 7.

As a simple example to illustrate the modelling process, consider the following problem. A London-based managing director whose company has a factory near the centre of Manchester, is woken early one morning by a telephone call from the Manchester factory. There has been a major industrial dispute during the night shift and in order to prevent a total walk out of the day staff, and a consequent closing of the production process (which would be very costly), he is required at the factory in Manchester as soon as possible. His problem is to get from his home in London to the factory in Manchester in minimum time. His possibilities are to travel by car, train or plane. So his problem is to choose the form of transport that will get him to his destination as quickly as possible.

We now move onto box 2, and formulate the model. Each type of journey can be divided into three parts, the time from his home to the starting point of the transport, including the waiting time, the time on the particular form of transport and the time required from the transport's stopping place to the final destination. To put this into mathematical terms, let T_i denote the time required for the ith transport where

$i = 1$ refers to car
$i = 2$ refers to train
$i = 3$ refers to plane.

Thus

$$T_i = a_i + b_i + c_i$$

where

a_i = time to get from initial point to start of transport i plus waiting time

b_i = time on transport i

c_i = time from stopping point of transport i to final destination.

In terms of our original diagram we are now moving onto box 3. Before we do so, it is important to note that we are setting up a general frame work which will solve not only this particular managing director's problem but any similar problem.

On to box 3. Here we state the problem in mathematical terms. We require to find the value of i (1, 2 or 3) which minimises the function T_i, and we can immediately move onto box 4 and write the solution as

$$j, \text{ where } T_j = \min (T_1, T_2, T_3)^*.$$

So our mathematical world is soon dealt with and requires little sophistication. We just need to compare the values T_1, T_2, T_3 and choose the minimum one. The next box is straightforward as well. The interpretation of solution j is that the transport labelled j should be used and we now apply this solution to the managing director's problem.

The map below illustrates the geographical problem. The managing director lives just outside London at Romford in Essex, and the factory is very close to the centre of Manchester. We now estimate each function T_i.

i = 1: CAR

Clearly for this form of transport $a_1 = c_1 = 0$, and we just need to evaluate b_1, the actual journey time in the car.

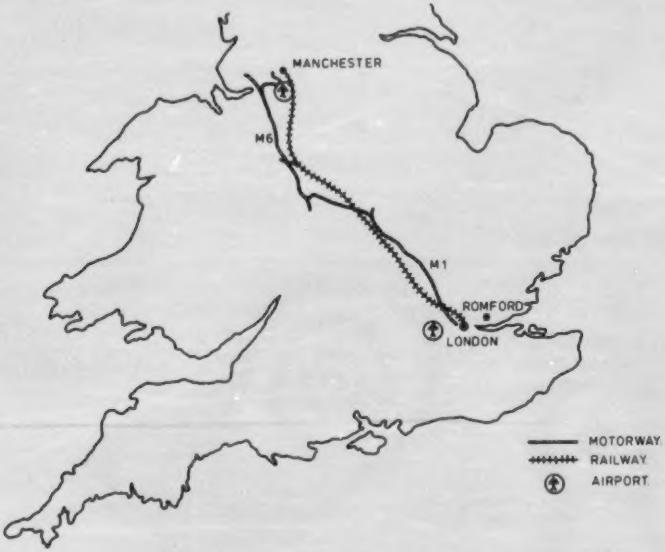


Fig. 2.

There is first a journey on poor quality roads across to the M1, which he estimates to take 70 minutes. This is followed by 170 miles of driving at 70 m.p.h., which will take

$$\frac{170}{70} \text{ hours} \simeq 146 \text{ mins.}$$

Allowing for a 20 minute journey from the motorway into Manchester and a 20 minute driving break (for breakfast!) we obtain

$$b_1 = 70 + 146 + 20 + 20 = 256$$

i.e. $T_1 = 256$ mins.

$i = 2$: TRAIN

The first part of this journey is by car to Euston Station and he estimates this as 70 mins, allowing for a possible 20 mins wait at the station for the train to Manchester; i.e. $a_2 = 70$. The journey time is 155 mins, i.e. $b_2 = 155$; and the taxi journey from Manchester Piccadilly Station to the factory takes 10 mins, i.e. $c_2 = 10$. Thus

$$T_2 = 70 + 155 + 10 = 235 \text{ mins.}$$

$i = 3$: PLANE

The journey across London to Heathrow by car is estimated to take 150 mins, allowing 30 mins to board the plane; i.e. $a_3 = 150$. The plane's time to Manchester Ringway Airport is 40 mins, i.e. $b_3 = 40$; and the final journey into the factory is estimated to take 30 mins; i.e. $c_3 = 30$. Thus

$$T_3 = 150 + 40 + 20 = 220 \text{ mins.}$$

The solution is given by j where

$$T_j = \min (T_1, T_2, T_3).$$

Using the above values, the answer is $j = 3$; i.e. travel by plane. So we have completed box 5 with the conclusion "travel by plane".

The manager now considers carefully whether the model allows for *all* features that might influence his ability to make an early intervention in the dispute (box 6). Experience on travelling on this route has shown him that there is a higher risk of planes being delayed and diverted and so ariving far later than their scheduled times than in the case of trains. He has found the extra delay quite frequently to be well in excess of the 15 minutes which is *theoretically* saved by travelling by plane instead of train (i.e. $T_2 - T_3$).

The risk that delay on the plane would be so great that he would fail to arrive in time to avoid the production process being closed down, makes him decide to go by train.

What has the manager done in the language of the mathematical modelling process? He has added to the model the important feature which we call *uncertainty*. The value $T_2 = 235$ mins is unchanged because in his experience trains on this route run punctually. A modified assumption in box 2 is made, however, to allow for a *spread of possible values* of T_3 from 220 mins to some much higher values. Then, a modified box 3 asks for the minimum journey time that can be *relied upon*. The revised answer (box 4) is $j = 2$; the interpretation

---

*Here the notation $x_j = \min (x_1, x_2, \ldots, x_n)$ means that x_j is the minimum value of $x_1, x_2, \ldots, x_n$.

(box 5) is "travel by train"; and the manager, now confident that the model has taken into account all relevant considerations (box 6), takes this decision (box 7).

The above example has illustrated the modelling process. We are not advising that it particularly helps using this formal process to solve *this* problem, but it does show clearly the various stages of modelling. We can fill in the relevant boxes for this particular problem, moving round the modelling cycle twice.

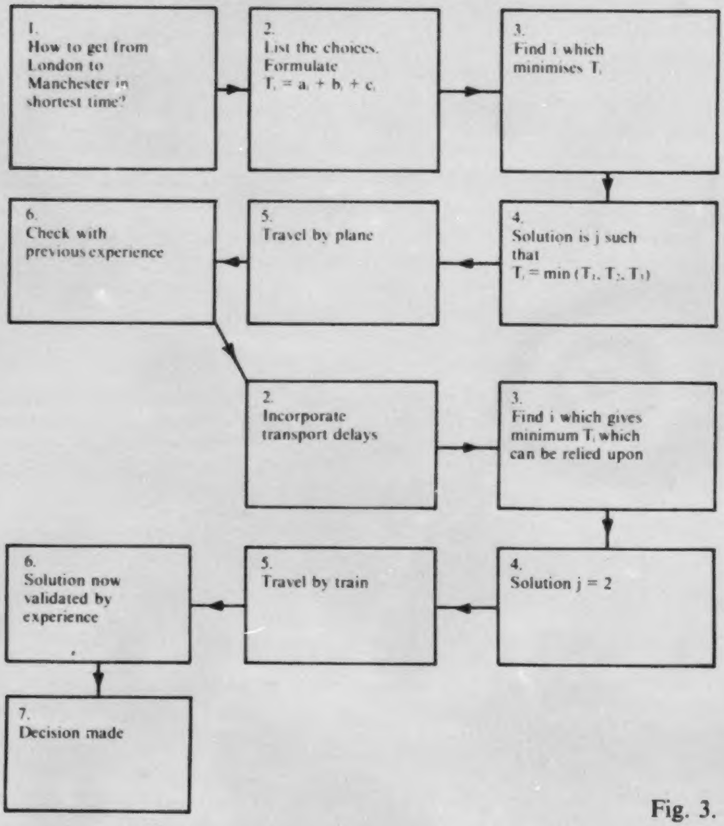

1. How to get from London to Manchester in shortest time?
2. List the choices. Formulate $T_i = a_i + b_i + c_i$
3. Find i which minimises T_i
4. Solution is j such that $T_j = \min(T_1, T_2, T_3)$
5. Travel by plane
6. Check with previous experience
2. Incorporate transport delays
3. Find i which gives minimum T_i which can be relied upon
4. Solution j = 2
5. Travel by train
6. Solution now validated by experience
7. Decision made

Fig. 3.

In the examples which follow, it might well be beneficial for you to use the boxes to illustrate the various stages of the modelling process. We must emphasise though that not all modelling will follow the exact pattern as shown. It is just a guide to what modelling is about.

2 – Forecasting: The Retail Price Index

In a period of high inflation it is evident that prices of most goods tend to rise. The actual rate at which prices increase varies a good deal. During the severe German inflation of 1920-23, prices increased by more than a thousand million times. In Britain, the Government publishes each month the *retail price index*, which is calculated by choosing a range of commodities that at a particular time cost £100. The same or equivalent commodities are then costed each month. The list is supposed to be typical for an average family and would include food, heating, rates, housing and travel, etc. The official table below shows how the price index has changed over a period of years, taking the January 1974 cost as £100. So that if Mr. and Mrs. Average spent £100 on their shopping in January 1974, then to buy the same goods in January 1964 would have cost them £54.6, and to buy them in January 1976 would cost £147.9.

	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1963	53.5	54.0	54.1	54.2	54.2	54.2	53.9	53.7	53.9	54.1	54.2	54.3
1964	54.6	54.6	54.8	55.3	55.8	56.0	56.0	56.2	56.2	56.3	56.7	56.9
1965	57.1	57.1	57.3	58.4	58.6	58.8	58.8	58.9	58.9	59.0	59.2	59.5
1966	59.6	59.7	59.7	60.5	60.9	61.1	60.8	61.2	61.1	61.2	61.6	61.7
1967	61.8	61.8	61.8	62.3	62.3	62.5	62.2	62.0	61.9	62.4	62.8	63.2
1968	63.4	63.7	63.9	65.1	65.1	65.4	65.4	65.5	65.6	65.9	66.1	66.9
1969	67.3	67.7	67.9	68.7	68.6	68.9	68.9	68.7	68.9	69.5	69.6	70.1
1970	70.6	71.0	71.4	72.5	72.7	72.9	73.5	73.4	73.8	74.6	75.1	75.6
1971	76.6	77.1	77.7	79.4	79.9	80.5	80.9	81.0	81.1	81.5	82.0	82.4
1972	82.9	83.3	83.6	84.4	84.8	85.4	85.6	86.3	86.8	88.0	88.3	88.7
1973	89.3	89.9	90.4	92.1	92.8	93.3	93.7	94.0	94.8	96.7	97.4	98.1
1974	100.0	101.7	102.6	106.1	107.6	108.7	109.7	109.8	111.0	113.2	115.2	116.9
1975	119.9	121.9	124.3	129.1	134.5	137.1	138.5	139.3	140.5	142.5	144.2	146.0
1976	147.9	149.8	150.6	153.5								

Fig. 4. Table – Retail Price Index

This table is about the past; the problem we now consider is that of estimating what price changes are likely in the future. It is of the utmost importance for government planning to be able to forecast the likely level of price increases in the coming months. It is also vital to management and union leaders in their negotiations on wage settlements. We perhaps ought to note here how the *rate of inflation* is actually evaluated. The yearly rate of inflation at April 1976, say, is found by evaluating the increase in the retail price index over the past year and then expressing this as a percentage of the retail price index at April 1975. Thus

$$\text{yearly inflation rate at April 1976} = \frac{(153.5 - 129.1)}{129.1} \times 100$$
$$= 18.9\%.$$

This is not the same quantity as the monthly inflation rate. At April 1976, the monthly rate is given by

$$\text{monthly inflation rate at April 1976} = \frac{(153.5 - 150.6)}{150.6} \times 100$$
$$= 1.9\%$$

The yearly rate (which works out at an average 1.6% per month) is again talking about the past, and the monthly rate at April 1976 is probably more significant if we are attempting to predict the likely future rate.

We return to our central problem; namely that of predicting the future retail price index. The simplest model used for forecasting is to assume that in the short run, the underlying value is constant but subject to random variations about the mean. So, for example, we might choose to average the previous six monthly values of the index in order to give an estimate of the mean, and then use this as our forecast. Applying the model to the data above in order to estimate the July 1963 figure we obtain the following calculation.

Month	Index	Total in last 6 months	6 month moving average	Forecast	Actual Value
Jan. 1963	53.5				
Feb. 1963	54.0				
Mar. 1963	54.1				
Apr. 1963	54.2				
May 1963	54.2				
June 1963	54.2	324.2	54.0	54.0	53.9
July 1963	53.9	324.6	54.1	54.1	53.7
Aug. 1963	53.7	324.3	54.05	54.05	53.9
Sept. 1963	53.9	324.1	54.0	54.0	54.1
Oct. 1963	54.1	324.0	54.0	54.0	54.2
Nov. 1963	54.2	324.0	54.0	54.0	54.3
Dec. 1963	54.3	324.1	54.0	54.0	54.6
Jan. 1964	54.6	324.8	54.1	54.1	54.6

Using the average of the previous six readings gives a forecast of 54.0. The actual value was 53.9, which gives good agreement between the model and reality. Continuing in this way, always using the previous six monthly values as the average (called the 6 monthly "moving average") gives the values in the fourth column which form the forecast in the fifth column. The actual values are given in the final column.

Analysing the figures in the last two columns shows that if there is any trend, then the forecast will always lag well behind. From July 1963 onwards the index is slowly rising whereas the *forecast* does not start to increase until December 1963. So this model is certainly of little use if there is a definite trend in the figures. This fact is even more apparent if we use the model at April 1976 in order to predict the May 1976 value. The 6 monthly moving average at April 1976 is

$$(144.2 + 146.0 + 147.9 + 149.8 + 150.6 + 153.5)/6 = 148.7$$

So the forecast for May 1976 is 148.7. This is hardly likely to be realised, since the April value was 153.5.

We have been round our modelling cycle once already, but we are clearly not happy with the model we have been using; so we must retrace our steps and go round the cycle once again. What is our main criticism of the "moving average" model? It does not take into account any trend in the figures. Looking at the retail price index, particularly from 1974 onwards, there is a definite increasing trend. If for example we use a 5 month moving average on the data, this is an estimate for the underlying value (i.e. without random variations) of the index at the midpoint of the third month. We can continue to calculate the 5 month moving average, but also we can estimate the trend by evaluating the differences between successive moving averages. If the increase is a linear one, we would expect the trend to be constant. So if we assume that the trend is subject to random variations, it is probably advisable to take a 3 month moving average of the trend. These calculations have been evaluated for the retail price index from May 1975 to April 1976 as shown below. (M.A. stands for moving average.)

Month	Index	5 month M.A.	Trend	3 month M.A. of the trend
1975 May	134.5			
June	137.1			
July	138.5	138.0		
Aug.	139.5	139.6	+ 1.6	
Sept.	140.5	141.0	+ 1.4	1.50
Oct.	142.5	142.5	+ 1.5	1.53
Nov.	144.2	144.2	+ 1.7	1.70
Dec.	146.0	146.1	+ 1.9	1.73
1976 Jan.	147.9	147.7	+ 1.6	1.80
Feb.	149.8	149.6	+ 1.9	
Mar.	150.6			
April	153.5			

At the end of April, the best estimate of the trend is 1.80. In order to produce the estimate for May 1976, the *moving average* must be projected forward. The latest figure is 149.6 which is the underlying figure at Feb. 1976. So we must project this forward three months to give us an estimate for May 1976. So our forecast for May 1976 is

$$149.6 + (3 \times 1.80) = 155.0$$

The actual value was 155.2, so we have reasonable agreement between the model and reality for this one prediction. Continuing to the next month, the next 5 month M.A. is 151.4 with trend 1.8 and 3 month M.A. of the trend 1.77. Thus our June 1976 prediction is

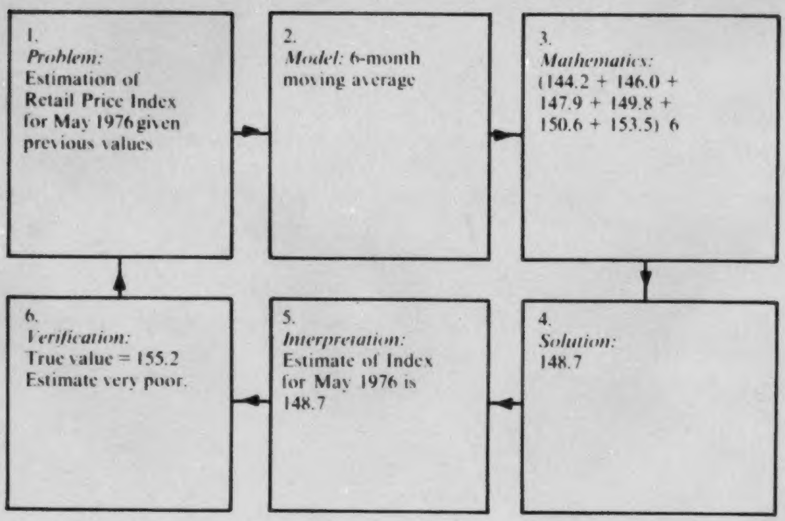
$$151.4 + (3 \times 1.77) = 156.7$$

and the actual value is 156.0. Again agreement between model and reality is reasonable. Using this same method, the predictions and actual results for the rest of 1976 are given below.

Month	Index	
	Predicted	Actual
1976 May	155.0	155.2
June	156.7	156.0
July	158.3	156.3
August	159.0	158.5
September	160.4	160.6
October	161.6	163.5
November	163.7	165.8
December	165.9	168.0

The correlation between the predicted and actual index values is fairly good until the end of the year. It over-estimates the July figure, but the actual July figure is probably somewhat less than would be expected, perhaps due to seasonal factors. Towards the end of the year though, the index starts to rise very rapidly, and the predictions are well short of the actual values. This in fact illustrates how prices were out of control towards the end of 1976. The model is based on a *linear* increase in the index – in practice the actual increases were more rapid for a time.

We have completed our modelling cycle once more, and this time we are somewhat happier with the model, although it still clearly has limitations. There are many other methods of estimation, the best known being exponential smoothing. It is worthwhile retracing our modelling steps in this example, using the boxes introduced in the first section:


1.
Problem: Estimation of Retail Price Index for May 1976 given previous values
2.
Model: 6-month moving average
3.
Mathematics: (144.2 + 146.0 + 147.9 + 149.8 + 150.6 + 153.5)/6
4.
Solution: 148.7
5.
Interpretation: Estimate of Index for May 1976 is 148.7
6.
Verification: True value = 155.2 Estimate very poor.

We return to box 1 since the first model is of little use.

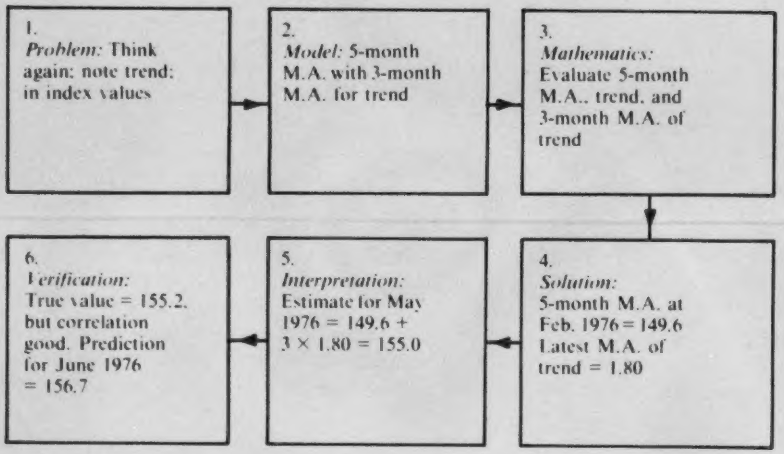

1.
Problem: Think again: note trend: in index values
2.
Model: 5-month M.A. with 3-month M.A. for trend
3.
Mathematics: Evaluate 5-month M.A., trend, and 3-month M.A. of trend
4.
Solution: 5-month M.A. at Feb. 1976 = 149.6 Latest M.A. of trend = 1.80
5.
Interpretation: Estimate for May 1976 = 149.6 + 3 × 1.80 = 155.0
6.
Verification: True value = 155.2, but correlation good. Prediction for June 1976 = 156.7

Fig. 5.

Further verification of the model shows that in periods of rapidly increasing inflation, an even more sophisticated model is required.

It is also worthwhile to note that the models described here are based solely on the data values given and the trends in such data. We call models of this type *empirical models*. The retail price index is clearly affected by a whole network of government, private and international decisions, and in order to make estimates a year ahead for example, a model incorporating these features is needed. The treasury has a very sophisticated model for predicting future economic trends, but many of the relevant decisions are very difficult to predict. For example, what level of wage increases will be negotiated nationally for the coming year, or what will happen to the foreign exchange rate? So even using very complicated and detailed models for the economy, predictions must be treated with caution.

Related Problems

1) The R.P.I. for 1977 and 1978 are given below

1977	172.4	174.1	175.8	180.3	181.7	183.6	183.8	184.7	185.7	186.5	187.4	188.4
1978	189.5	190.6	191.8	194.6	197.7	197.2	198.1	199.4	200.2	201.1	202.5	204.2

Use the model above, or design your own, in order to estimate the next *five* values (Actual values are given at the end of this booklet).

2) Design a model to forcast R.P.I. which takes into account the seasonal fluctuations.

3) The F.T. Index on the London Stock Market for the first two weeks in June 1979 is given below. Estimate its value for the next two weeks. (Actual values are given at the end of this booklet).

510.9 509.3 516.9 514.0 503.2
503.9 501.4 488.9 474.2 478.5

3 – Handicapping: Weightlifting

Weightlifting is of comparatively modern origin, the first world championship being staged in London in 1891. Most competitions comprise two types of lifts, namely the snatch and the jerk. For the snatch lift, the weight is pulled from the floor to a locked arms overhead position in a single move, although the lifter is allowed to move or squat under the weight as it is being lifted. For the jerk lift, two movements are allowed, the first to the chest and the second to the overhead position.

At the moment, there are nine official bodyweight classes as given in Table 1. The problem we consider here is how to compare lifters in different classes. Is it possible to design a handicap system so that all competitors can compete together on equal terms?

Bodyweight class	Maximum weight in class (kg)
Flyweight	52
Bantamweight	56
Featherweight	60
Lightweight	67.5
Middleweight	75
Light-heavyweight	82.5
Middle-heavyweight	90
Heavyweight	110
Super-heavyweight	no limit

In order to design a fair handicapping system we must first formulate a mathematical model to show how theoretically the lift varies with the weight of the lifter. Clearly the lift will increase with increasing bodyweight, but there is no reason to suppose for instance that there should be a linear relationship. We must use the model to determine the relationship.

In setting up a mathematical model the initial stage is to identify the important variables. In this case we define

L – the maximum lift, and
W – the bodyweight of the lifter.

The first assumption in the model is to suppose that L is proportional to the average cross-sectional area, say A, of the lifter's muscle. In mathematical terms we express this as

$$L = k_1 A, \tag{1}$$

where k_1 is a constant. The second assumption is to suppose that A is proportional to a 'typical' body length, say l, squared. Expressing this in mathematical terms,

$$A = k_2 l^2 \tag{2}$$

where k_2 is a constant. The 'typical' body length can be defined from

$$W = k_3 l^3 \tag{3}$$

where k_3 is the average body density. We also make the assumption that constants k_1, k_2 and k_3 are the same for all lifters. This means that for all competitors

$$L = k_1 A = k_1 k_2 l^2 = k_1 k_2 (W/k_3)^{2/3}$$

and writing $K = k_1 k_2 / k_3^{2/3}$,

$$\boxed{L = KW^{2/3}} \tag{4}$$

So the model predicts that $L/W^{2/3}$ should be a constant for all lifters. Before attempting to use $L/W^{2/3}$ as the handicapped lift, the model must be checked with known data to make sure that it is a reasonable one. For example we can use the 1976 Montreal Olympic Games results for both the snatch and jerk lifts. In Table 2, we list the winning lifts, L_s for snatch, L_j for jerk, for each class division, and using the upper limit in each weight class as W, the quantities $K_s = L_s/W^{2/3}$ and $K_j = L_j/W^{2/3}$ are evaluated.

Weight Class	W	L_s	L_J	$L_S/W^{2/3}$	$L_J/W^{2/3}$
Flyweight	52	105.1	137.7	7.55	9.89
Bantamweight	56	117.7	145.2	8.04	9.92
Featherweight	60	125.2	160.2	8.17	10.45
Lightweight	67.5	135.2	172.8	8.16	10.43
Middleweight	75	145.2	190.3	8.17	10.70
Light-heavyweight	82.5	162.7	202.9	8.59	10.71
Middle-heavyweight	90	170.3	212.8	8.48	10.60
Heavyweight	110	175.3	225.4	7.64	9.82

Table 2 — 1976 Montreal Olympic Games Results

The last two columns show that K_S and K_J for each weight class are roughly constant. The average value for K_S is 8.10 and for K_J is 10.31. If we accept the model as a justifiable basis for handicapping, we can use the ratio $L/W^{2/3}$ as the handicapped lift; the winning competitor is the one with the largest value of $L/W^{2/3}$. Returning to the 1976 Olympic Games, the overall champion for both the snatch and lift is the light-heavyweight champion, who was Z. Kaczmarek of Poland.

The model is of course a very simple one. A closer look at the validation of it with the Olympic Games data perhaps indicates a bias towards the lifters in the middle range. On the other hand, lifters in these classes are probably not overweight or underweight and so might well be expected to produce better handicapped lifts.

We can again use the box notation introduced in §1 to describe the modelling process here.

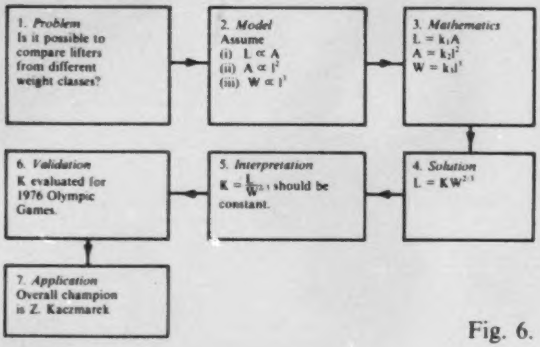

1. *Problem*
Is it possible to compare lifters from different weight classes?
2. *Model*
Assume
(i) $L \propto A$
(ii) $A \propto l^2$
(iii) $W \propto l^3$
3. *Mathematics*
$L = k_1 A$
$A = k_2 l^2$
$W = k_3 l^3$
4. *Solution*
$L = KW^{2/3}$
5. *Interpretation*
$K = \frac{L}{W^{2/3}}$ should be constant.
6. *Validation*
K evaluated for 1976 Olympic Games.
7. *Application*
Overall champion is Z. Kaczmarek

Fig. 6.

Related Problems

1. The model for handicapping weightlifting on BBC Television's "Superstars" programme is to take L-W as the handicapped lift. Comment on its suitability for (i) handicapping serious weightlifting competitors
   (ii) its use on the TV programme.
2. Another method of obtaining power-law models is to take a log-log plot of the data. If, for example,
   $$L = KW^{\alpha}$$
   then $\log L = \log K + \alpha \log W$. This means that a $\log L - \log W$ graph should be a straight line, with slope α. Use the data given above to estimate the most appropriate value of α.
3. The winning times in the 1976 Olympic Games running events are given below. Use a log-log plot of the data to produce a power-law relationship between distance and time.

Distance (m)	Time (s)
100	10.06
200	20.23
400	44.26
800	103.50
1500	219.17
5000	804.76
10000	1660.38

4 – Planning: Train Time Tables

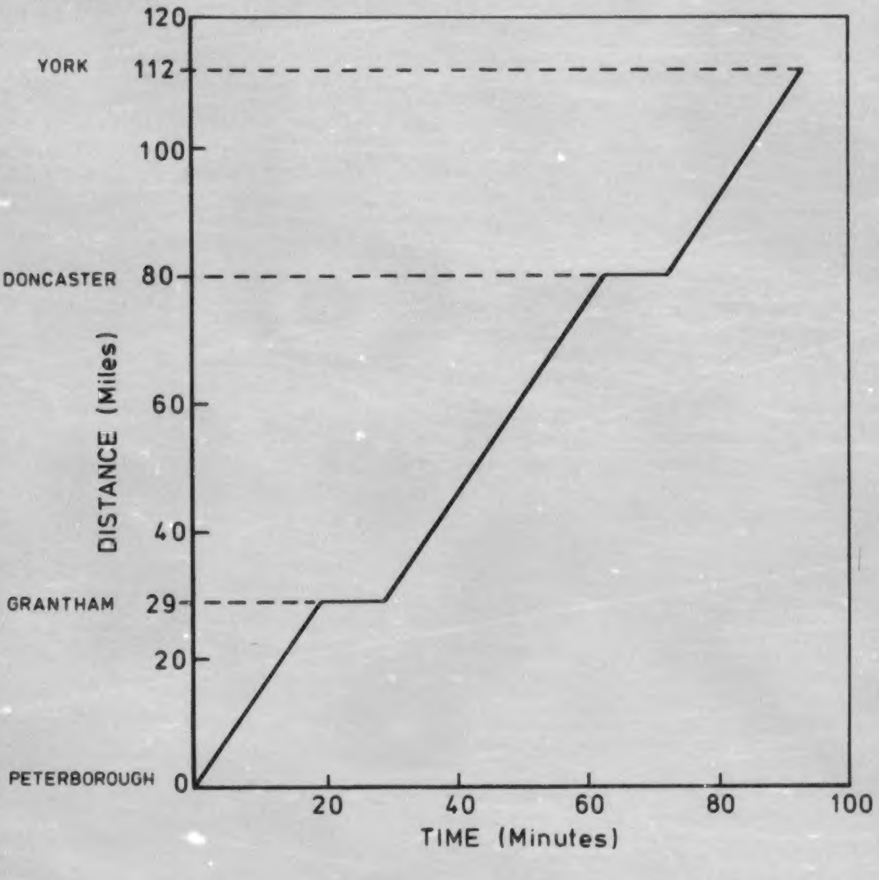
One of the difficult problems faced by British Rail is the construction of viable time tables for sections of single track which are used by trains of different speeds and where passing can only take place at intermediate stations. As an example, take the introduction of H.S.T. (high speed trains) on the East Coast main line between Peterborough and York. H.S.T. travel at an average speed of about 120 m.p.h. whereas the conventional expresses cruise at about 90 m.p.h. Suppose we have a normal express train travelling from Peterborough to York with intermediate stops at Grantham and Doncaster, and we wish to time table H.S.T. which run non-stop from Peterborough to York passing the normal express at one of the intermediate stations. How many minutes should separate the express departures from Peterborough?

A method of illustrating the problem is to show the time table in graphical form. Figure 7 illustrates the history of the normal express assuming that

(i) between stations it runs at constant speed 90 m.p.h.

(ii) it stops at each intermediate station for 10 minutes.

Of course both assumptions are only approximations to reality. The train has to accelerate to its cruising speed and to decelerate to zero speed at each intermediate station. Also the track probably does not permit cruising at 90 m.p.h. throughout its length. Consequently the stopping time of 10 minutes at intermediate stations is an over-estimate of that required in reality. In this way some of the limitations of the first assumption are compensated in the second assumption.



120
YORK
112
100
DONCASTER
80
DISTANCE (Miles)
60
40
GRANTHAM
29
20
PETERBOROUGH
0
20
40
60
80
100
TIME (Minutes)

Fig. 7.

In figure 8 we illustrate the graph of an H.S.T. using the same assumption (i) as above, running non-stop from Peterborough to York.

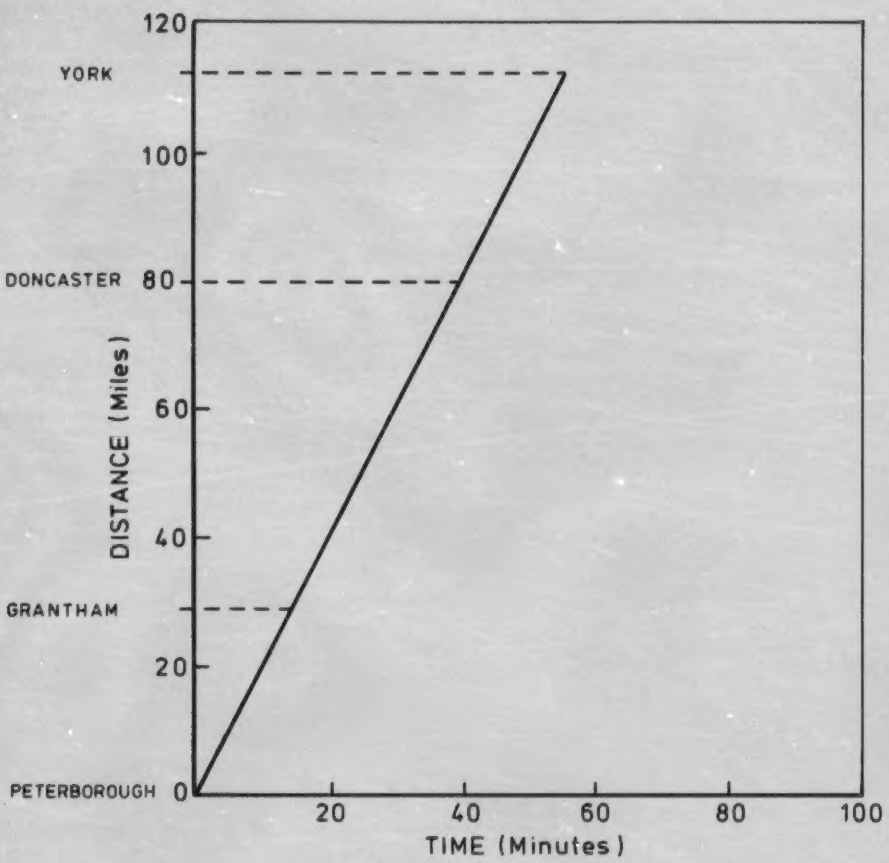


Fig. 8.

Solutions to the problem can be found by superimposing the H.S.T. graph on the express train graph, the overtaking occurring when the express train is stationary at the intermediate stations or at the final station. The graph shows that there are three possible solutions, corresponding to the H.S.T. leaving Peterborough at 10, 27 and 42 minutes after the express train. This is illustrated in figure 9.

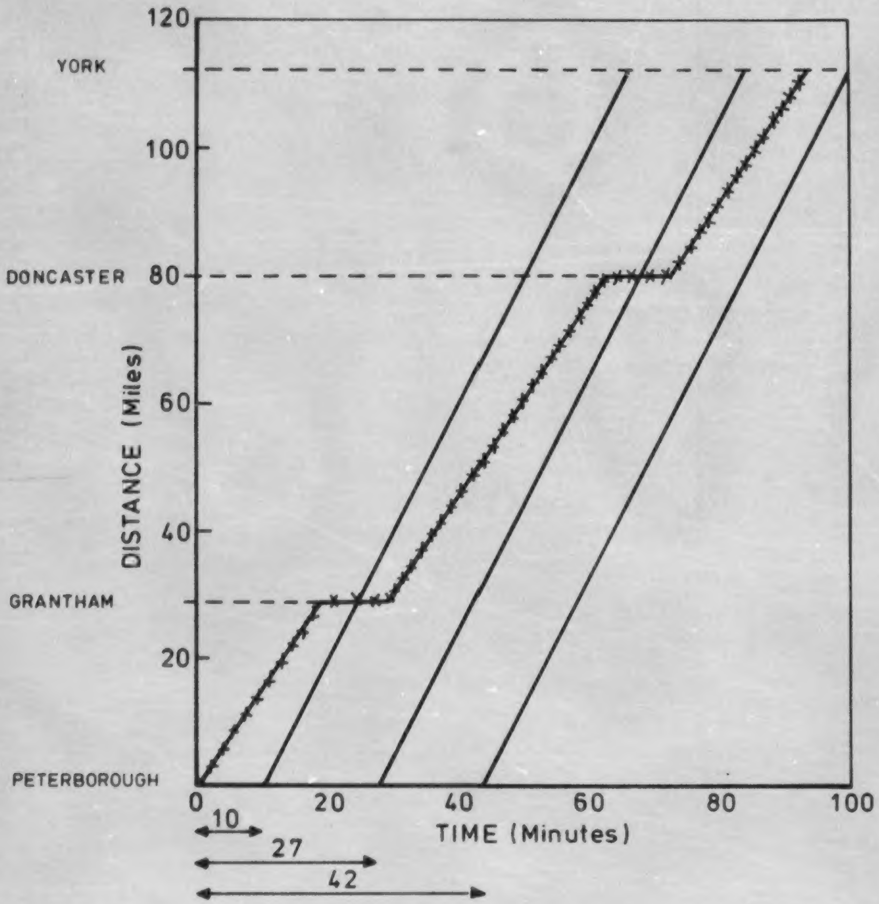


Fig. 9.

Notice that by scheduling the H.S.T. to pass the express train at an instant just halfway through the express train's stopping time we give maximum allowance for possibilities of timing errors. For example, if passing were scheduled near the beginning of the express train's scheduled stopping time, a very small delay in the express train's journey would force the signalling system to bring the H.S.T. to a halt. Again, if passing were scheduled near the end of the scheduled stopping time, any delay in the journey of the H.S.T. would force the express train to depart late.

Of course we can readily check our graphical answers by translating the model into mathematical terms. First define

D_1 = distance travelled by normal express train
D_2 = distance travelled by H.S.T.

Then, with reference to figure 1 and using distance = speed × time gives

$$D_1 = \begin{cases} \frac{3}{2}t & 0 \leqslant t \leqslant t_1 \\ 29 & t_1 \leqslant t \leqslant t_1 + 10 \\ \frac{3}{2}(t - 10) & t_1 + 10 \leqslant t \leqslant t_2 \\ 80 & t_2 \leqslant t \leqslant t_2 + 10 \\ \frac{3}{2}(t - 20) & t_2 + 10 \leqslant t \leqslant t_3 \end{cases}$$

(t is measured in minutes and D_1 in miles)

and t_1, t_2, t_3 are given by

$$\frac{3}{2}t_1 = 29, \qquad \frac{3}{2}(t_2 - 10) = 80, \quad \frac{3}{2}(t_3 - 20) = 112,$$

i.e. $t_1 = 58/3$, $t_2 = 190/3$, $t_3 = 284/3$.

If τ denotes the time gap at Peterborough between the two trains then

$$D_2 = 2(t - \tau)$$

and the possible choices for τ are given by

(i) $D_2 = 29$ when $t = t_1 + 5$
(ii) $D_2 = 80$ when $t = t_2 + 5$
(iii) $D_2 = 112$ when $t = t_3 + 5$

These in turn give $\tau = 9.8, 28.3, 43.7$ minutes.

So our graphical solution gave a reasonably accurate prediction; in fact quite accurate enough for the model used. The next stage would be to improve the model by

(i) allowing for acceleration and braking of the trains,
(ii) putting into the model hazards along the line such as speed restrictions at, for example, Rossington due to danger of subsidence in an extensive mining area and at Selby where the track has a very sharp bend.

Related Problems

1. Advanced Passenger Trains (A.P.T.) are soon to be introduced on the London Euston line to Liverpool, Manchester and Glasgow. These trains have a *maximum* speed of 150 m.p.h. whereas conventional expresses on these routes have a *maximum* speed of 100 m.p.h. Suggest how to incorporate 5 of these trains into the timetable.

   (At present, there is a basic hourly service to Liverpool, an hourly service to Manchester and a two hourly service to Glasgow).

2. Using your suggested service pattern from the problem above, how many platforms would be needed at Euston station to accommodate the trains?

5 – Predicting: Population Changes

Accurate population predictions are clearly needed by both local and regional government. Local authorities need to plan the allocation of their resources. They must make decisions about investment in their housing programme, or in the provision of social services, or in the education service. If new buildings are required these normally take many years to plan and build, so decisions often have to be taken on the basis of estimations of future levels of population. It is just as vital for central government to have accurate predictions for the population. An increase in the birth rate would for instance require a larger number of nursery and primary teachers in years to come. It takes at least three years to train a teacher, so a gradual increase in the training of teachers would be required in order to cope with the increase in school children which would occur in a few years time. On the other hand a continual decrease in the birth rate (which we have seen in the U.K. for more than 10 years now) also requires important decisions to be made. Fewer teachers would be required, so recruitment must fall; existing teachers might need to be redeployed in subjects which are still short of teachers, and buildings would be available for other uses. All the decisions that need to be taken are very much dependent on future levels of the birth rate.

Many mathematical models have been used in an attempt to accurately predict the population changes. We will start our analysis by considering a very simple model. If this does not give adequate predictions we will try to modify and improve the model. Our problem then is to predict the future population of a country. We first identify the important variables, which in this case are population, say x, and time. Suppose $x_1, x_2, x_3, \ldots, x_i$ denote the population level at the beginning of each year, so that the change in population during one year is

$$x_{i+1} - x_i.$$

This increase (or decrease) is due to births and deaths (assuming no migration or emigration), so that

$$x_{i+1} - x_i = \text{births} - \text{deaths}.$$

157

The most simple way of modelling births and deaths is to suppose that both are proportional to the current population level, i.e.

$$\begin{cases} \text{births} = bx_i \\ \text{deaths} = dx_i \end{cases}$$

where b and d are positive constants. With these assumptions

$$x_{i+1} - x_i = bx_i - dx_i$$

which gives

$$x_{i+1} = (1 + b - d)x_i.$$

Writing $c = 1 + b - d$,

$$\boxed{x_{i+1} = cx_i} \qquad (i = 0, 1, 2, \ldots) \qquad (1).$$

This is a difference equation but it is easy to solve. Clearly

$$x_1 = cx_0$$

and

$$x_2 = cx_1 = c(cx_0) = c^2x_0.$$

In the same way,

$$x_3 = cx_2 = c(c^2x_0) = c^3x_0.$$

and it should now be clear that

$$\boxed{x_i = c^ix_0} \qquad (i = 0, 1, 2, \ldots) \qquad (2).$$

This is the solution to the model, but it is of little use unless we have values for c and x_0. For example if x_0 and x_1 are known,

$$x_1 = cx_0$$

giving $c = (x_1/x_0)$. Then for $i = 2, 3, 4, \ldots$

$$x_i = \left(\frac{x_1}{x_0}\right)^i x_0 \qquad (3).$$

We must now attempt to validate the model. Population statistics are well documented and so we have numerous examples on which to test the model. As one example, we consider the population of England and Wales from 1880 onwards. The actual values for x_0 and x_1, which represent the populations at the beginning of 1880 and 1881 respectively are

$$x_0 = 25{,}714 \times 10^3, \quad x_1 = 26{,}046 \times 10^3.$$

Thus $c = 1.0129$ and we can use (3) to predict the population from 1881 onwards.

These results are given below, together with the actual population.

Predicted Year	Actual Population ($\times 10^3$)	Population ($\times 10^3$)	Difference ($\times 10^3$)
1880	25,714	25,714	0.0
1881	26,046	26,046	0.0
1882	26,382	26,334	0.048
1883	26,723	26,627	0.096
1884	27,067	26,922	0.145
1885	27,417	27,220	0.197
1886	27,770	27,522	0.248
1887	28,128	27,827	0.301

The two sets of figures do show some correlation, but the last column which gives the difference in population between the predicted and real values, clearly shows that the differences are increasing. In fact a closer examination of these differences shows a very regular increase indeed. This implies that one method of obtaining an improved model, given that we are now back at 1881, would be to use a better value for c, perhaps one obtained not by just using the 1880 and 1881 population figures but also taking into account earlier population data. This brings us back to the methods of forecasting, encountered in 2. We will leave this approach to the reader. For information, earlier population data is given in the table below.

Year	Population ($\times 10^3$)
1875	24,045
1876	24,370
1877	24,700
1878	25,033
1879	25,371

Our new approach is to return to the model itself and decide if any of the assumptions made were unreasonable and if so to try and rectify the situation. If we look back at the formulation, there are only two basic assumptions, namely

(i) births and deaths are proportional to the population size
(ii) no migration and emigration.

It is relatively easy to take into account the second factor, but the first assumption is more fundamental, and rather suspect. The actual number of births and deaths must be influenced not just by the overall population size but by the age structure of that population. The age structure of a population of a country can have wide variations. For example the age structure of the U.K. and Mauritius for 1959 are compared in figure 10. In forecasting future trends it is clear that the age structure together with birth and death rates for each age division must be accounted for.

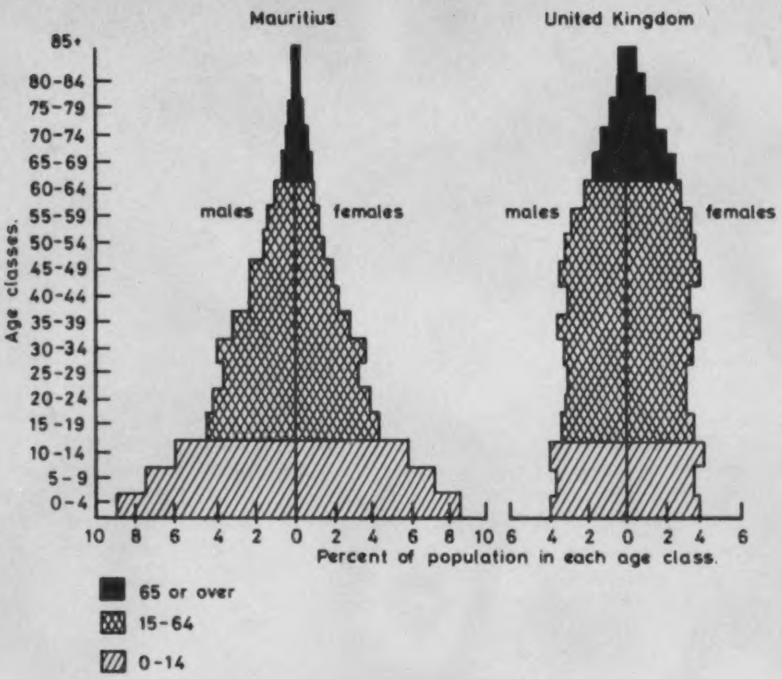


Fig. 10.

Rather than use the 17 age divisions indicated in figure 10, let us attempt a simpler model with just three age groups, namely 0-14, 15-44, 40+. Denote the populations in each of these groups by $n_1(t)$, $n_2(t)$, $n_3(t)$, t referring to the year. It is convenient to introduce a population matrix (or vector):

$$N(t) = \begin{bmatrix} n_1(t) \\ n_2(t) \\ n_3(t) \end{bmatrix}$$

Let b_i and d_i denote the birth and death rates for each age group. We have chosen the age ranges so that $b_1 = b_3 = 0$ (approximately). Consequently the population in group 1 after one year is made up of the number born together with those of group 1 who have survived the year but not advanced into group 2. Assuming an even distribution in the age group (which is not true of Mauritius in 1959), this second number is given by $\frac{14}{15}(1 - d_1)n_1$. Hence

$$n_1(t + 1) = b_2 n_2(t) + \frac{14}{15}(1 - d_1)\, n_1.$$

Similarly

$$n_2(t+1) = \frac{1}{15}(1-d_1)n_1 + \frac{29}{30}(1-d_2)n_2$$

$$n_3(t+1) = \frac{1}{30}(1-d_2)n_2 + (1-d_3)n_3.$$

Again, it is far more convenient to represent these equations in matrix form, so that

$$\begin{bmatrix} n_1(t+1) \\ n_2(t+1) \\ n_3(t+1) \end{bmatrix} = \begin{bmatrix} \frac{14}{15}(1-d_1) & b_2 & 0 \\ \frac{1}{15}(1-d_1) & \frac{29}{30}(1-d_2) & 0 \\ 0 & \frac{1}{30}(1-d_2) & (1-d_3) \end{bmatrix} \begin{bmatrix} n_1(t) \\ n_2(t) \\ n_3(t) \end{bmatrix}$$

or

$$\boxed{N(t+1) = AN(t)} \qquad (4)$$

where

$$A = \begin{bmatrix} \frac{14}{15}(1-d_1) & b_2 & 0 \\ \frac{1}{15}(1-d_1) & \frac{29}{30}(1-d_2) & 0 \\ 0 & \frac{1}{30}(1-d_2) & (1-d_3) \end{bmatrix}$$

The 3 × 3 matrix A is referred to as the *transition* matrix. If the coefficients of A are known and the initial population matrix N(0) given, equation (4) can be repeatedly used to find N(1), N(2), N(3), . . . , etc. In fact

$$N(1) = AN(0)$$

and

$$N(2) = AN(1) = A(AN(0)) = A^2N(0).$$

Similarly $N(3) = A^3N(0)$, and continuing in this way

$$\boxed{N(i) = A^iN(0)} \qquad (5)$$

where i is a positive integer. Note the similarity between the matrix difference equation (4) and the solution (5), and the earlier scalar difference equation (1) and its solution (2). The earlier model is just a special case of this one.

We are now at the validation stage again, and we will still use the England and Wales population data for 1881. Estimates for the birth and death rates taken from the population figures up to 1881 are

$b_2 = 0.075$

$d_1 = 0.20$, $d_2 = 0.008$, $d_3 = 0.062$

and the age distribution into the three age groups in 1881 is given by

$$N(0) = \begin{bmatrix} 9{,}498 \\ 11{,}641 \\ 4{,}907 \end{bmatrix}$$

Using 1000 people as a unit for the population figures given, the transition matrix A is given by

$$A = \begin{bmatrix} 0.915 & 0.075 & 0.0 \\ 0.065 & 0.959 & 0.0 \\ 0.0 & 0.033 & 0.938 \end{bmatrix}$$

Thus, from (4)

$$N(1) = AN(0) = \begin{bmatrix} 9{,}575 \\ 11{,}781 \\ 4{,}987 \end{bmatrix}$$

giving a total population size equal to $26{,}343 \times 10^3$. Continuing in this way,

$$N(2) = AN(1) = \begin{bmatrix} 9{,}645 \\ 11{,}920 \\ 5{,}067 \end{bmatrix}$$

and the total is $26{,}632 \times 10^3$;

$$N(3) = AN(2) = \begin{bmatrix} 9{,}719 \\ 12{,}058 \\ 5{,}146 \end{bmatrix}$$

with total $26{,}923 \times 10^3$;

$$N(4) = AN(3) = \begin{bmatrix} 9{,}797 \\ 12{,}195 \\ 5{,}225 \end{bmatrix}$$

with total $27{,}217 \times 10^3$;

$$N(5) = AN(4) = \begin{bmatrix} 9{,}879 \\ 12{,}332 \\ 5{,}303 \end{bmatrix}$$

with total $27{,}514 \times 10^3$;

$$N(6) = AN(5) = \begin{bmatrix} 9{,}964 \\ 12{,}469 \\ 5{,}381 \end{bmatrix}$$

with total $27{,}814 \times 10^3$.

The predicted and real values are compared in the table below.

Year	Predicted Population ($\times 10^3$)	Actual Population ($\times 10^3$)	Difference ($\times 10^3$)
1881	26,046	26,046	0
1882	26,343	26,334	9
1883	26,632	26,627	5
1884	26,923	26,922	1
1885	27,217	27,220	−3
1886	27,514	27,522	−8
1887	27,814	27,827	−13

The agreement between the predicted values and the true values is now very good. The predicted values for 1882 is only 0.03% too high, and this extremely small error is probably due to the fact that in 1881 the death rate for children in the age range 0-4 was smaller than usual.

Of course the high degree of accuracy for this model applied to the population of England and Wales in 1881 is dependent upon very steady values for the birth and death rates. If these values are changing appreciably even this model would have marked limitations, and would need to be modified again. If conditions remain fairly steady a population structure with 2 age groups (0-14, 15+) can give reasonable answers. Examples of this type are simpler to evaluate since they involve multiplying 2 × 2 and 2 × 1 matrices, and so are perhaps more suitable for pupils to work out themselves.

Related Problems

1) Develop a two age group model to predict population change, and use it to predict the England and Wales population for 1882 onwards using the data above.
2) Statistics for the United Kingdom population, 1966-1970, are given below.

Population (in thousands)

	1966	1967	1968	1969	1970
Persons: all ages	54,500	54,800	55,049	55,263	55,421
Under 5 years	4,790	4,791	4,759	4,696	4,586
Under 18 years	15,219	15,315	15,451	15,576	15,672
Under 21 years	17,815	17,963	17,943	17,962	17,992
5-14	8,047	8,212	8,407	8,590	8,781
15-29	11,258	11,323	11,401	11,474	11,568
30-44	10,241	10,107	10,012	9,924	9,843
45-64	13,493	13,541	13,525	13,494	13,436
65-74	4,269	4,365	4,454	4,554	4,637
75 and over	2,402	2,461	2,491	2,531	2,570

Birth Rates

	Total per 1000 Population	Total per 1000 women, aged 15-44
1966	18.0	91.7
1967	17.6	90.3
1968	17.2	89.1
1969	16.7	86.7
1970	16.3	85.2

Death Rates (per 1000 population)

Males

Ages	0	1-4	5-9	10-14	15-19	20-24	25-34	35-44	45-54	55-64	65-74	75-84	over 85
1966	12.4	5.3	0.5	0.4	1.0	1.1	1.1	2.5	7.4	21.8	54.0	121.6	261.0
1967	11.8	4.8	0.4	0.4	1.0	1.0	1.0	2.4	7.1	20.9	51.7	113.5	240.1
1968	12.4	4.9	0.4	0.4	0.9	0.9	1.0	2.4	7.2	21.4	54.1	123.2	265.5
1969	12.5	4.8	0.4	0.4	0.9	0.9	1.0	2.4	7.4	21.9	55.5	119.8	250.4
1970	12.3	4.8	0.4	0.4	0.9	1.0	1.0	2.3	7.2	21.4	54.0	119.0	253.4
Females													
1966	11.2	4.1	0.3	0.3	0.4	0.5	0.7	1.8	4.5	10.5	29.2	82.2	206.7
1967	10.6	3.9	0.3	0.3	0.4	0.4	0.6	1.7	4.3	10.2	27.6	76.0	192.8
1968	11.3	3.8	0.3	0.3	0.4	0.4	0.7	1.7	4.4	10.6	28.7	80.9	215.6
1969	11.3	3.7	0.3	0.2	0.4	0.4	0.6	1.8	4.5	10.8	29.1	78.1	203.2
1970	11.2	3.6	0.3	0.2	0.4	0.4	0.6	1.7	4.5	10.5	28.3	77.0	205.0

Develop a model to predict the population of United Kingdom for the next five years (Actual data is given at the end of this booklet).

6 – Estimating: Attractiveness of Competing Shopping Centres

One of the difficulties faced by planners of new shopping centres is to estimate the likely success of the venture. Will it produce sufficient customers for the retail shops to make a profit, how large should the centre be, how much car parking space should be provided, from how far away will it attract shoppers, etc.? Those of you who have been to the Brent Cross Shopping Centre in North London will realise that in peak times (particularly Saturdays) the parking facilities are far from adequate. One explanation for this is that this centre attracts shoppers from a far wider area than expected. A significant number of shoppers have travelled further than 20 miles in order to shop at Brent Cross. The problem we will discuss here is how to estimate the attractiveness of competing shopping centres.

The basic variable we shall work with is T, the average number of trips made (in some time interval) by individuals living at a distance D from the shopping centre. Consider a residential zone, distance D from the shopping centre, Clearly as D increases, T will decrease, fairly steeply at first but less dramatically as D continues to increase. We can model this sort of behaviour by assuming that

$$\boxed{T = \frac{KF}{D^2}} \tag{1}$$

where K is a constant, and F is a measure of the attractiveness of the centre. Its value must be estimated by taking account of such factors as

(i) how many shops are there in the centre; (sometimes the retail floor space is used)

(ii) how many departmental stores are there; is there a branch of Marks and Spencer in the centre;

(iii) how many supermarkets are there;

(iv) what other attractions are there in the centre (e.g. libraries, museums, cinemas, parks, etc.)?

For example consider the neighbouring towns of Northampton and Wellingborough. Northampton, serving a local population of over 100,000, has a large shopping centre which includes a Debenham Departmental store and a branch of Marks and Spencer. It also includes a theatre, cinema, library, museum and art gallery. On the other hand, Wellingborough, which has a population of just over 30,000, has only a limited number of the amenities that Northampton enjoys. Its only real advantage is that it is probably easier to park near to the centre. So if F_1 denotes the degree of attractiveness of Northampton and F_2 for Wellingborough, we estimate that $F_2 = \frac{1}{3} F_1$. This means that we rate their attractiveness roughly in proportion to the populations which they support, and this is probably a reasonable way of comparing attractiveness of conventional *urban* shopping centres.

For planning problems it would be useful to estimate the area of influence of competing shopping centres. We want to know the boundary of choice between one shopping centre and another. On one side of the boundary, shoppers will tend to go to one centre more often; on the other side, they will tend to go more frequently to the other centre. So suppose we have a shopping centre A at (0, 0), using rectangular coordinates (x, y) and a neighbouring centre B at (a, 0) ($a > 0$). We consider a residential zone at P (x, y). The situation is illustrated in figure 11.

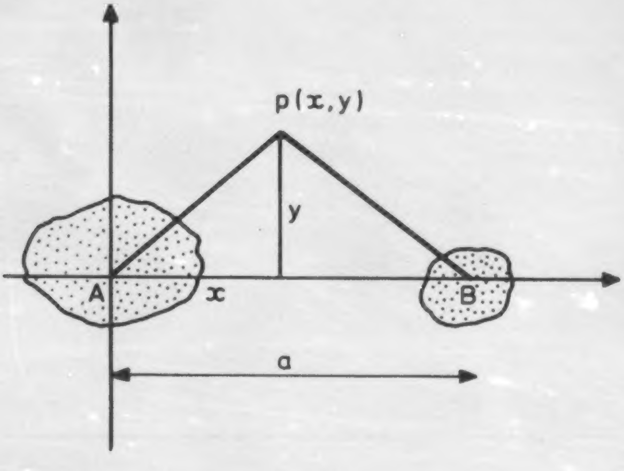


Fig. 11.

For points on the boundary of choice, the number of trips made to each centre will be equal and so if P (x, y) is on the boundary,

$$\frac{KF_1}{D_1^2} = \frac{KF_2}{D_2^2} \qquad (2).$$

But $D_1 = x^2 + y^2$ and $D_2 = (a - x)^2 + y^2$, giving

$$c\left[\,(a - x)^2 + y^2\,\right] = x^2 + y^2$$

where $c = F_1/F_2$. Thus

$$\frac{ca^2}{(c-1)} - \frac{2cax}{(c-1)} + x^2 + y^2 = 0$$

or

$$\left[x - \frac{ca^2}{(c-1)}\right] + y^2 = \frac{ca^2}{(c-1)^2} \qquad (3).$$

This is the equation of a circle centre $x = \frac{ca}{(c-1)}$, $y = 0$ and radius $ca/(c - 1)$.

This is illustrated in figure 12 for the case $c = 3$.

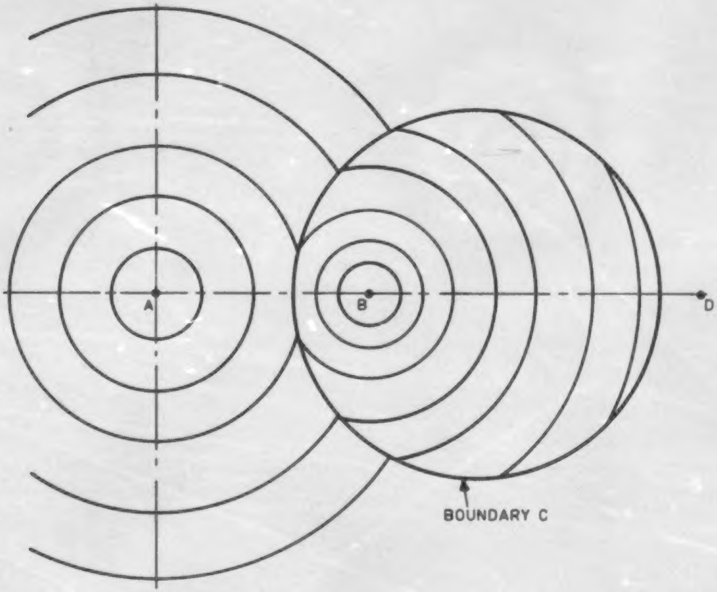


Fig. 12

The circles round each centre are curves on which T is constant. Note that for all points outside the boundary C, the shopping centre preferred is A. For example points to the far east of the boundary (e.g. point D) would prefer to bypass the centre at B and shop at the centre A, which is further away in distance but with more attractive amenities.

Let us apply this theoretical model to a practical example. We earlier introduced the relative attractiveness between the towns of Northampton and Wellingborough for which we estimated $c = F_1/F_2 = 3$. This is the case illustrated in figure 12, and superimposing this diagram on a map of the area around Northampton and Wellingborough results in figure 13.

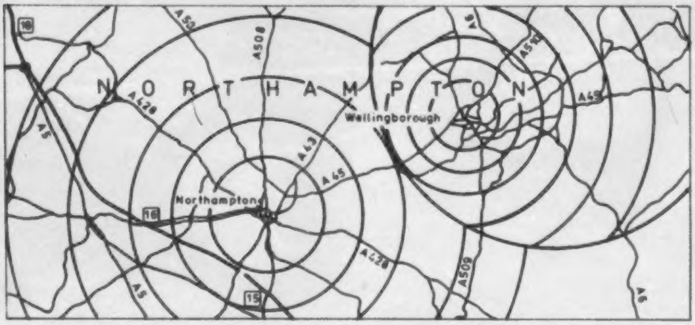


Fig. 13.

According to the model a number of villages such as Bozeat and Sywell, which are in fact closer in distance to Wellingborough, are principally attracted to Northampton. It would be interesting to try to validate the model by conducting surveys in villages which are near the boundary of attractiveness between the two centres. For inhabitants without cars, it will probably be found that the most significant factor is probably not the attractiveness of centres or distances but the availability and frequency of bus services into the shopping centres.

One of the uses of models of this type is the planning of new shopping centres. For example if it is planned to develop a new hypermarket together with easy parking, discount petrol, children's playground and many other attractions at a site between Northampton and Wellingborough, it is important to estimate the likely use of the proposed centre. If we rate the proposed centre as twice as attractive as Northampton (and so six times more attractive than Wellingborough) it is easy to determine the boundaries of choice between the three shopping centres. If the proposed site for the new centre is situated midway between the two towns, we have the situation illustrated in Figure 4. Points P on the boundary of choice between A and C satisfy equation (3) with $c = \frac{1}{2}$ and a replaced by $a/2$, i.e.

$$\left(x + \frac{a}{2}\right)^2 + y^2 = \frac{a^2}{2}$$

which is the circle of centre $x = -a/2$, $y = 0$ and radius $a/\sqrt{2}$. To determine the boundary of choice between C and B, we introduce new axes x^1, y^1 with C as the origin. We can again use equation (3) with $c = 6$ and a replaced by $a/2$ i.e.

$$(x^1 - \tfrac{3}{5}a)^2 + (y^1)^2 = \frac{3a^2}{50}$$

which is a circle centre $x^1 = \frac{3}{5}a$ (which means $x = \frac{11}{10}a$) and radius $\sqrt{\frac{3}{50}}\,a$.

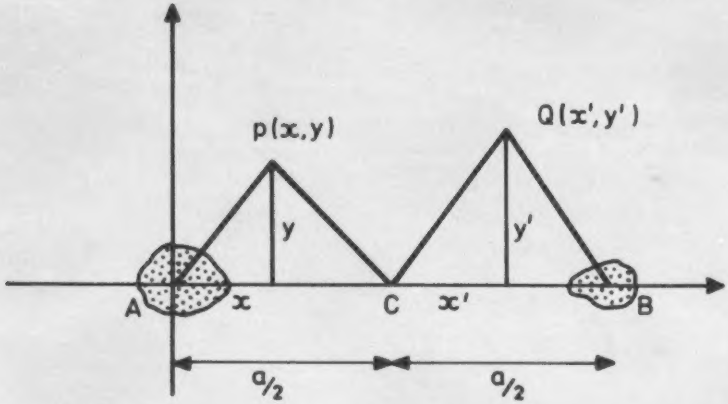

p(x,y)
Q(x',y')
y
y'
A
x
C
x'
B
a/2
a/2

Fig. 14.

The two boundaries are illustrated in figure 15 on the Northampton-Wellingborough map. While not providing actual predictions about the number of shoppers, it does clearly show the devastating effect that the proposed new centre would have on the shops in Wellingborough. It predicts that people in such places as Wollaston, Rushden and even Irthingborough would have preference for the proposed new centre. To actually predict the likely number of shoppers requires a more sophisticated model.

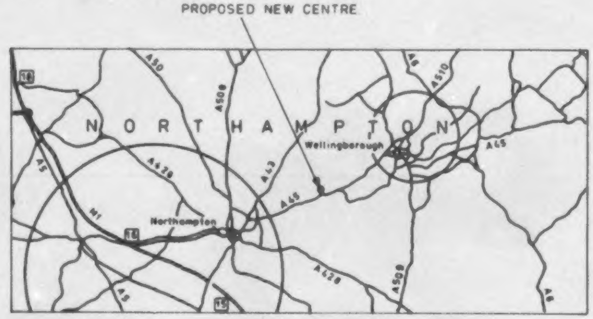

PROPOSED NEW CENTRE
A50
A508
A6
A510
N O R T H A M P T O N
Wellingborough
A45
A428
A43
A5
M1
Northampton
16
15
A45
A428
A509
A5
A6

Fig. 15

Returning to the introductory remarks made about the Brent Cross Shopping Centre, it is of interest to estimate the number of people arriving at the Shopping

Centre from distance D (in a prescribed time interval). We can approximate the *area* which is at distance D miles (to the nearest mile) from the Shopping Centre by $(2\pi D) \times 1 = 2\pi D$ square miles.

If N is the population per square mile, then the number of people living in this area (see figure 6) is $2\pi DN$ and the number of trips by these people is

$$(2\pi DN)\, T = (2\pi DN)\left(\frac{KF}{D^2}\right) = \frac{2\pi NKF}{D} \qquad (4)$$

The interesting thing to notice about (4) is that the number of trips is inversly proportional to D; which falls off far less dramatically than D^2. Unfortunately, although models of this type are used by planners, as yet we have had little attempt at verification of such models.

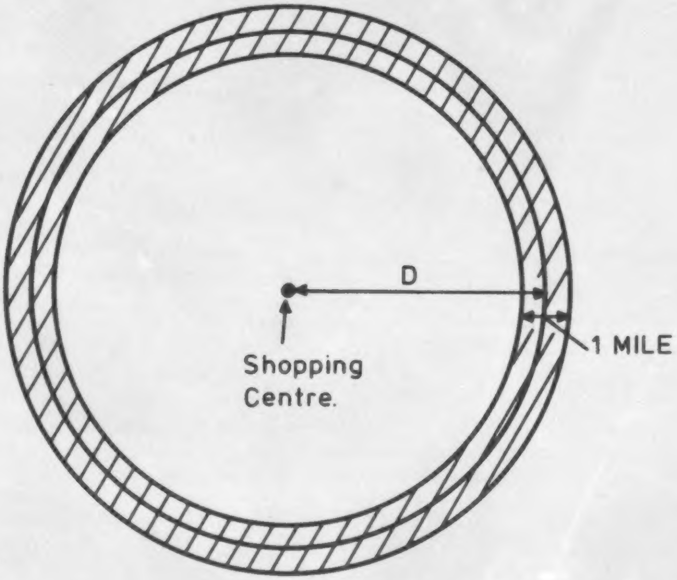


Fig. 16.

Related Problems

1). Develop a model to measure the attractiveness of Shopping Centres. Apply to your nearby Shopping Centre and see if the measure gives reasonable results. If not, improve the model.

2). If the form taken from T is

$$T = \frac{KF}{D^{\propto}}$$

where $\propto$ is a positive number, develop the model along the lines outlined above.

7 – Decision Making: Stocking Magazines

In this section we will illustrate some of the ways in which decisions are made. We will describe the models using a simple stock decision problem.

Suppose a newsagent has to decide how many copies of a specialist motoring magazine to stock each week. Demand for the magazine fluctuates, sometimes as many as four are sold each week whereas at other times none at all are sold. In most weeks he will sell 2 or 3. The magazine costs the newsagent 60p and its selling price is £1.00, but no refunds are available from the publisher for unsold copies. The newsagent's problem is to decide what stock to order each week.

Assuming that he will never sell more than 4 copies the newsagent has five possible decisions:

Stock 0 magazines
Stock 1 magazine
Stock 2 magazines
Stock 3 magazines
Stock 4 magazines.

These are also 5 possible outcomes, namely:

0 magazines demanded
1 magazine demanded
2 magazines demanded
3 magazines demanded
4 magazines demanded.

The pay-off table for the various possible decisions and outcomes is illustrated in a matrix formulation below.

		Decision				
		0	1	2	3	4
Outcome	0	0	-0.60	-1.20	-1.80	-2.40
	1	0	0.40	-0.20	-0.80	-1.40
	2	0	0.40	0.80	0.20	-0.40
	3	0	0.40	0.80	1.20	0.60
	4	0	0.40	0.80	1.20	1.60

When outcomes can be predicted with complete certainty, it is usually easy to determine the optimum decision; but as in this example if uncertainty is present, no single decision will be best for all outcomes.

If the newsagent is a born pessimist he would assume the worst was going to happen each week and would therefore choose to stock 0 magazines. On the other hand if he is an optimist he would stock 4 magazines. These two decision criteria are called

maximin (i.e.,maximise the minimum possible pay-off)
maximax (i.e. maximise the maximum possible pay-off)

Another decision criterion is called "minimax regret", which means minimise the maximum extent to which you may regret that you did not take another decision after the event. For example, if the newsagent stocks one magazine, and only one is required, he has no regrets. But if two are required his regret is £0.40, the profit he could have made on the second magazine. If three magazines are required, his regret is £0.80 and so on. Continuing in this way we can determine his regret for all the possible outcomes if he stocks 1 magazine.

Number Required	0	1	2	3	4
Regret (£)	0.60	0	0.40	0.80	1.20

So the newsagents maximum possible regret if he stocks one magazine is £1.20. We can analyse the situation for each decision and in each case determine the maximum regret.

Decision	0	1	2	3	4
Maximum Regret (£)	1.60	1.20	1.20	1.80	2.40

So using the *minimax regret* criterion, his decision would be to stock either 1 or 2 magazines.

The basic criticism of the decision criteria introduced above is that they do not take into account how likely the various possible outcomes are. In order to have a reasonable basis for decision making, we must have some way to estimate the probabilities of occurrence of each of the outcomes. Suppose the newsagent had kept a record of past demand for the magazine over the last 50 weeks. This is shown below.

Demand	No. of weeks it occurred	Probability
0	5	0.10
1	12	0.24
2	18	0.36
3	11	0.22
4	4	0.08

The probabilities are just found by dividing the number of weeks when a certain demand occurred by 50, the total number of weeks observed.

Suppose now that 4 magazines are stocked. The expected profit per week to be made is determined by multiplying the profit made for each outcome by the probability of its occurrence and summing all the terms, i.e.

$$\begin{aligned}\text{expected profit} = £\ & (0.10)(-2.40) + (0.24)(-1.40) + (0.36)(-0.40) \\ & + (0.22)(0.60) + (0.08)(1.60) \\ & = -£0.46\end{aligned}$$

The reason why we calculate the expected profit in this way can be seen by considering 100 typical weeks. The total profit can be expected to be -£2.40 in 10 weeks, -£1.40 in 24 weeks, -£0.40 in 36 weeks, £0.60 in 22 weeks, £1.60 in 8 weeks. This gives a total profit of -£46 over a hundred weeks, so that the average expected profit per week is -£0.46.

We also calculate the expected weekly profit for the other courses of action. Stocking 3 magazines leads to an expected profit

$$\begin{aligned}£\ & (0.10)(-1.80) + (0.24)(-0.80) + (0.36)(0.20) + (0.22)(1.20) \\ & + (0.08)(1.20) \\ & = £0.06;\end{aligned}$$

stocking 2 magazines leads to an expected profit

$$\begin{aligned}£\ & (0.10)(-1.20) + (0.24)(-0.20) + (0.36)(0.80) + (0.22)(0.80) \\ & + (0.08)(0.80) \\ & = £0.36;\end{aligned}$$

stocking 1 magazine leads to an expected profit

$$\begin{aligned}£\ & (0.10)(-0.60) + (0.24)(0.40) + (0.36)(0.40) + (0.22)(0.40) \\ & + (0.08)(0.40) \\ & = £0.30;\end{aligned}$$

and finally stocking zero magazines clearly leads to zero profit. The table below summarises the results.

Number stocked	Expected Weekly Profit (£)
0	0
1	0.30
2	0.36
3	0.06
4	-0.46

The optimum decision now emerges as

"stock 2 magazines each week".

If the occurrence of the various demands continues with the same frequency, the expected average profit will be £0.36 per week, which is higher than the average profit from taking any other decisions.

The decision criterion used is called "maximise expected monetary value" i.e. maximise EMV; and it has wide applicability to decision situations which occur over and over again.

Related Problem

If the past selling record is given in the table below, what decision should be taken.

Demand	No. of weeks it occurred
0	7
1	11
2	17
3	12
4	5

8 – Modelling in the Classroom

In the preceding sections we have given examples of the use of mathematics for forecasting, handicapping, planning, predicting, estimating and decision making. Our intention has been to show that mathematics does have a vital role to play in solving problems related to real situations. Many people in industry and commerce use mathematics as an integral part of their job. It is a key tool for so many people as well as being a beautiful subject in its own right.

This brings us to the difficult question of how to teach modelling in the classroom. Reading through an example on the application of mathematics is not the same as presenting that example to a class. It is important that pupils not only see how mathematics is used but also participate in the process of modelling. To achieve this requires a mature understanding between the class and the teacher. The teacher must attempt to lead the class in an appropriate direction but not by dictating. Ideas coming from the class should not be dismissed out of hand, simply because they are not the ones that the teacher had prepared. Remember that in modelling there is not a unique right answer. Some models can easily be discarded as being inadequate whilst for others it is a matter of critical judgement whether they present adequate solutions for the problem. For example in population modelling we have already seen that an age structure model gives good agreement with reality. Another type of population model can be formed using calculus and this can also be a very effective model in many situations. So modelling in the classroom requires both patience and understanding and a willingness to move in an unprepared direction. This clearly means the teacher needs confidence and an ability to know when to stop a particular line of action. Spending an hour getting nowhere can be frustrating for both the class and the teacher, although the time is not necessarily completely wasted, particularly if the class has played a participating role.

It should be clear from these remarks that we do not regard modelling as a substitute for normal mathematics teaching, particularly when a class is *not* enjoying mathematics. Quite the opposite in fact. If a class has a good working relationship with their mathematics teacher it should then be possible to spend some time on modelling. If the teaching is effective, pupils should enjoy the challenge of modelling and appreciate the role that mathematics plays in the outside world. As well as enjoying the participation in the modelling process, it will increase their enthusiasm for routine mathematical problems. Although we feel it important to relate mathematics to reality, we also appreciate it is not easy to achieve in the classroom and great care and patience must be employed.

We have already suggested that it is important for the class to participate but how far should this be taken? Should pupils be expected to actually go through the modelling process themselves? We feel that it is *not* reasonable and probably counter effective to expect pupils to model in isolation. The process of setting up a model, which is one of the major difficulties of the modelling process, is best done as a group activity; this also applies to the formation of the mathematical problem. The solution stage could be done on an individual basis, but the validation and application of the resulting solutions is again a difficult stage, where judgement and skill is needed, and this is most effectively achieved with close leadership. Only the brightest of your pupils will be able to model a new situation without great help. On the other hand, it is possible to set modelling problems similar to ones already considered, and these should help your pupils to gain confidence in their ability to apply mathematics to real problems.

We have deliberately not written out the examples in a form directly applicable to the classroom. In teaching modelling, so much depends on the teacher's individual style and personality, and so we do not seek to dictate the way in which a particular model should be taught. It is up to the teacher to prepare the material in the way which he or she feels is suitable. In the last section we provide references and extra material for the examples in the preceding sections. We hope that this information will help teachers in their preparation of the material. We have listed a number of books where other mathematical models are presented.

Opportunities for cooperation through 'modelling' are numerous. For example, several of the case studies in this booklet arise directly from the 'New Geography' and from Biology. *All* the sciences, however, provide numerous examples of 'physical' or 'real' problems for which a modelling approach is ideal. Why not talk about this with your science (or mathematics) colleagues? There is a great potential for broadening the scope of mathematics with many genuine science problems, and for applying mathematics more in the sciences. This potential can only be realised if individual teachers in schools get together to discuss this cooperation and actually cooperate.

Finally, for those about to embark on a modelling exercise in the classroom, good luck. Do not despair, and try to keep calm, even if events are not going in the direction you planned. It is not easy to teach mathematical modelling and you will find that it will take time and experience to teach material of this sort effectively. Nevertheless we feel it is worth persevering. The possible achievements are numerous. Your pupils will see the relevance of mathematical topics, they will learn to think constructively, they will gain insight into critical judgement and, given time, patience and understanding, they will gain confidence in their own ability to solve problems.

9 – References, Extra Material and Data for Related Problems

1. *Forecasting: The Retail Price Index*

A very readable account of elementary forecasting techniques is given in

(1) J. Hull, J. Mapes and B. Wheeler 'Model Building Techniques for Management' Saxon House (1976).

The techniques described in §2 and in (1) can of course be used for forecasting a variety of quantities, e.g. prices of individual goods, unemployment figures, population statistics, number of strikes, crime rates, stock market prices. Details of such statistics can be found in a variety of sources. One useful reference is the statistical section in the business supplement of the Sunday Times. Other sources are

(2) B. Mitchell 'Abstract of British Historical Statistics' C.U.P. (1971).

(3) A. Sillitoe 'Britain in Figures' Pelican (1971).

(4) Whittaker's Almanac.

2. *Handicapping: Weightlifting*

The model described here has been taken from

(5) M. S. Borrie and D. N. Burghes 'Mathematical Model for Weightlifting' Mathematics in School. Vol. 7 No. 1, 14.

There have been a number of recent articles relating mathematical models to sport, but most use fairly sophisticated mathematics. One article that deals with the use of mathematics in a variety of sports is

(6) J. Keller 'Optimal Velocity in a Race' American Mathematical Monthly Vol. 81, 474.

3. *Planning: Train Time Tables*

Details of train times and distances between stations are given in the British Rail Timetable, Published annually.

4. *Predicting*: *Population Changes*

Statistics for this section can be found in (2). There are many interesting mathematical books which have sections on populations; for example

(7) R. Haberman 'Mathematical Models' Prentice-Hall (1977).

(8) M. Holt and D. T. E. Marjoram 'Mathematics in a Changing World' Heinemann (1973).

(9) J. Lighthill 'Newer Users of Mathematics' Penguin (1978).

5. *Estimating*: *Attractiveness of Competing Shopping Centres*

This model is described in

(10) Unit 5 "Modelling by Mathematics" O.U. course TM281, together with other similar models; for example an agricultural land use model.

6. *Decision Making*: *Stocking Magazines*

This example is taken from (1). Other books on operational research with interesting examples are

(11) M. Wagner 'Principles of Management Science' Prentice Hall (1975).

(12) A. Battersby 'Mathematics in Management' Penguin (1962).

There are now a number of books on mathematical modelling. Unfortunately most of the models described either use fairly sophisticated mathematics or are not actually modelling a real situation.

(7), (8) and (9) all describe interesting models for a variety of situations. The Open University Course TM281 "Modelling by Mathematics" describes mathematical techniques which are important for technology and science. It is assumed that the reader has little previous knowledge of mathematics and it illustrates many interesting applications. Other books which include many mathematical models are

(13) S. Grossman and J. Turner 'Mathematics for the Biological Sciences' Macmillan (1974).

(14) M. Braun 'Differential Equations and their Applications' Springer Verlag (1975).

(15) P. Lancaster 'Mathematical Models of the Real World' Prentice hall (1976).

(16) T. McDonald 'Mathematical Methods for Social and Managerial Scientists' Houghton-Mifflin (1974).

(17) E. A. Bender 'An Introduction to Mathematical Modelling' Wiley (1978).

(18) M. L. Bettinger 'Calculus: A modelling approach' Addison-Wesley (1976).

A gentle introduction to the modelling process with some interesting examples is given in

(19) O.U. course M101 "Mathematics Foundation Course Block VI.

Data for Related Problems

(p 13) 1) 208.1, 208.9, 210.6, 214.2.

3) 482.9, 486.6, 485.1, 474.9, 476.0
473.4, 475.2, 466.9, 469.3, 473.4.

(p 29)
5.

	1971	1972	1973	1974	1975	1976
All ages	55,610	55,793	55,933	55,965	55,943	55,928
Under 5 years	4,503	4,405	4,295	4,122	3,952	3,739
Under 18 years	. .	. .	. .	. .	15,608	15,432
Under 21 years	. .	. .	. .	. .	17,956	17,859
5-14	8,908	9,023	9,088	9,128	9,120	9,099
15-29	11,714	11,870	11,974	12,054	12,150	12,302
30-44	9,751	9,709	9,765	9,824	9,865	9,925
45-64	13,390	13,316	13,222	13,120	13,028	12,936
65-74	4,736	4,820	4,910	4,988	5,043	5,079
75 and over	2,069	2,650	2,681	2,729	2,785	2,848

and Mathematics Teachers

ICSU Committee on the Teaching of Science

Edited by Alan Rogerson

⑥

Mathematical Modelling with Calculus

D. N. Burghes
and
M. S. Borrie

∫ydx
-b±√b²-4ac / 2a
0101
0001
≥
÷
×
C₂H₅OH
E=mc²
NaOH

Although this booklet is primarily aimed at the mathematical teacher, it should also be of interest to all teachers of science. It sets out a number of lively and varied case studies suitable for modelling treatment: traffic flow, road accidents, stock control, population changes and so on. Each situation is explored in detail and each illustrates the basic modelling flow-chart set out at the beginning of the booklet. Modelling is often a new way of looking at the *applications* of mathematics in the outside world, and in the sciences. We hope it will, therefore provide common ground for cooperation between mathematicians and scientists.

Contents

1 – Introduction 3
2 – King Arthur's Round Table 8
3 – Traffic Flow 11
4 – Clunk-Click 16
5 – Optimum Order Quantity 19
6 – Single Species Population Models 22
7 – Launching a Satellite 27
8 – 'Save It' – Water Heating and Cooling 32
9 – Conclusions 37
10 – References 38

Mathematical Modelling with Calculus

D. N. Burghes

Cranfield Centre for Teacher Services,
Cranfield Institute of Technology

&

M. S. Borrie

Department of Mathematics,
Pilgrim School,
Bedford.

1 - Introduction

In this booklet we will be concerned with mathematical models which require the use of calculus in the solution stage of the modelling process. We first introduce you to the concepts involved in mathematical modelling.

The main stages in mathematical modelling are summarised in the diagram below. We are concerned with a problem in the real world. We might want to explain some observed data, or make some predictions, or take a decision. To achieve this we translate the real problem into a mathematical one by making a number of simplifying assumptions. Important variables must be identified, and the relationships between them postulated. This leads to a mathematical problem of some sort, which is solved for the relevant variables using appropriate mathematical techniques. Clues for solving are often provided by bearing in mind the practical meaning of the mathematics. The solutions must now be interpreted back in terms of the real problem. Attempts should be made to validate the model, that is to check that the theoretical solution is in good agreement with the observations from the real situation. If there is good correlation, then the model can be used either to give a theoretical explanation for the observed phenomena, or to predict further results, or to help in making decisions. On the other hand if the correlation between the theoretical and observed results is not adequate we must return to the assumptions made in the model and decide which needs modifying or what additions should be made. The cycle is then traversed once more to see if the new model gives an adequate description of the real problem.

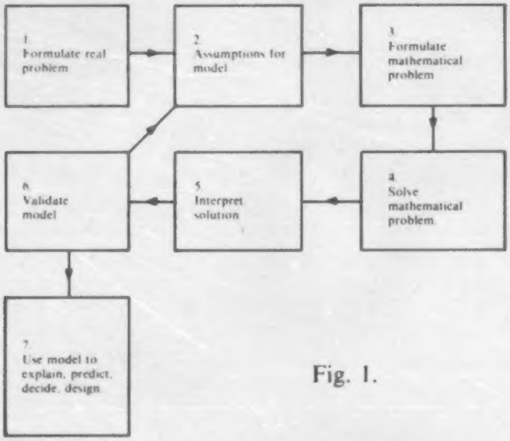

1
Formulate real problem
2
Assumptions for model
3
Formulate mathematical problem
4
Solve mathematical problem.
5
Interpret solution
6
Validate model
7
Use model to explain, predict, decide, design

Fig. 1.

The left hand column in the diagram represents the real world, the right hand column the mathematical world and the connections between these two worlds is in the middle column. Most mathematics teaching is solely concerned with box 4, although in applied mathematics there is sometimes a token movement through boxes 1, 2 and 3, but usually none back through boxes 5, 6 and 7. Unless these connections are given some importance in mathematics teaching, pupils will not realise that mathematics can and does play a vital role in solving many of today's problems in our scientific and industrial world.

It should be stressed at this stage that the modelling process as described above does not always take the precise form shown. The boxes are there just to give some idea of the underlying relationship between real world problems and the mathematical techniques used to find solutions to them. The important concepts to appreciate are the two translation stages, firstly from a real problem into a mathematical one through the model and secondly back from the mathematical solution to its interpretation in terms of the real problem. In order to illustrate these ideas we take a simple example.

Consider the problem of the motion of a smooth spherical mass which is lying on a smooth horizontal table, joined to one end of a spring which is fixed at the other end. The system is illustrated in figure 2, and we will concern ourselves with motions of the mass along the line of the spring.

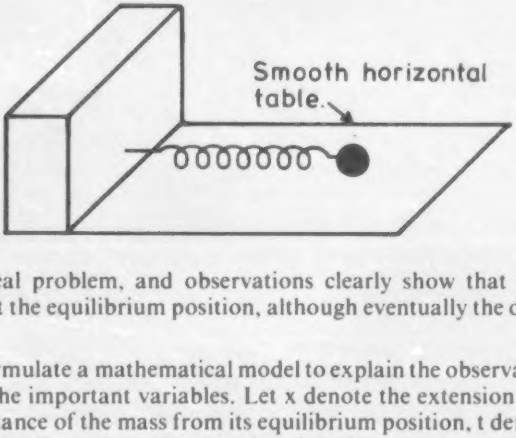


Fig. 2.

This is the real problem, and observations clearly show that the mass will oscillate about the equilibrium position, although eventually the oscillations die out.

In order to formulate a mathematical model to explain the observation, we must first identify the important variables. Let x denote the extension of the spring, and so the distance of the mass from its equilibrium position, t denote time, and T the tension in the spring. Clearly x is a function of t, i.e. $x = x(t)$, whilst T depends on the extension, i.e. $T = T(x)$. We make the following assumptions about the interaction of these variables.

(i) the tension in the spring is proportional to its extension (usually known as Hooke's Law);

(ii) the only force acting on the mass in the direction of motion is the tension (so we are ignoring any other forces, such as resistance to motion);

(iii) the force acting on the mass is equal to the rate of change of its linear momentum (this is Newton's second law of motion).

In mathematical terms, we can express these assumptions as

(i) $T = \lambda x$ (λ is a constant for a particular spring);

(ii) $F = -T$ if F denotes the total force in the direction of motion;

(iii) $F = \frac{d}{dt}(m \frac{dx}{dt}) = m \frac{d^2x}{dt^2}$ (since m is constant).

Eliminating T and F results in the differential equation

$$\frac{d^2x}{dt^2} + \left(\frac{\lambda}{m}\right) x = o \ . \qquad (1)$$

Writing $w^2 = \frac{\lambda}{m}$, the general solution of this equation is

$$x = A \cos (wt) + B \sin (wt) \ .$$

If initially, $t = 0$, $x = a$ say, and the mass is released, so that $\frac{dx}{dt} = 0$ at $t = 0$, then the appropriate solution is

$$x = a \cos (wt) \ . \qquad (2)$$

We now interpret this solution back in terms of the original problem. We first note that the displacement is predicted as oscillating. The solution is illustrated in figure 3. The period of the oscillation is $2\pi/w = 2\pi(m/\lambda)^{1/2}$, and the amplitude is a. So this model predicts that the mass performs oscillations about its equilibrium position with constant period and amplitude.

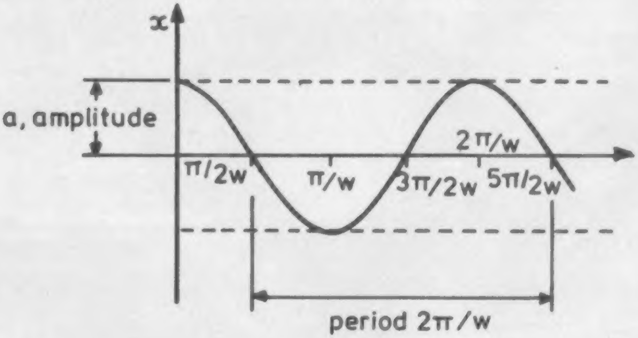


Fig. 3.

As a first attempt at a model, it is not unreasonable; but observations show that as time continues, the oscillations eventually die out. So the model does *not* explain all the observations. We must now re-examine our assumptions. Assumption (ii) seems the most likely to need modification. Although the mass and table are smooth, there will always be some friction between them as the mass moves. So suppose we modify (ii) to read

(ii) the only forces acting in the direction of motion are the tension in the spring and a resistive force, which we assume to be proportional to the speed of the mass.

In mathematical terms we now have

(i) $T = \lambda x$

(ii) $F = -T - mk\frac{dx}{dt}$ (the proportionality constant is taken as mk)

(iii) $F = m\frac{d^2x}{dt^2}$.

Hence
$$\frac{d^2x}{dt^2} + k\frac{dx}{dt} + w^2x = 0 \tag{3}$$

where $w^2 = \lambda/m$. The form of the solution of this differential equation depends on the values of the two parameters k and w; but assuming k to be small so that $k^2 < 4w^2$, we have the general solution

$$x = e^{-kt/2}\left[A\cos(\mu t) + B\sin(\mu t)\right]$$

where $\mu = (4w^2 - k^2)^{\frac{1}{2}}$. Again if at $t = 0$, $x = a$, $\frac{dx}{dt} = 0$, then the solution is

$$x = ae^{-kt/2}\cos(\mu t) \tag{4}$$

which is illustrated in figure 4.

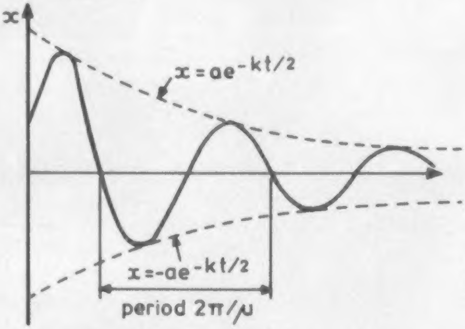


Fig. 4.

Oscillatory motion is still predicted, but this time the amplitude of the oscillations is decaying. This agrees with the observed behaviour and this model seems to provide a good explanation for all the observations. In order to use it to make predictions, estimates must be found from the experimental work for the values of λ and k.

It might appear that we have not covered any new ground in the above example but for us the important aspect of this example is to illustrate the modelling process. We can do this effectively using the boxes introduced earlier, and giving the following diagram.

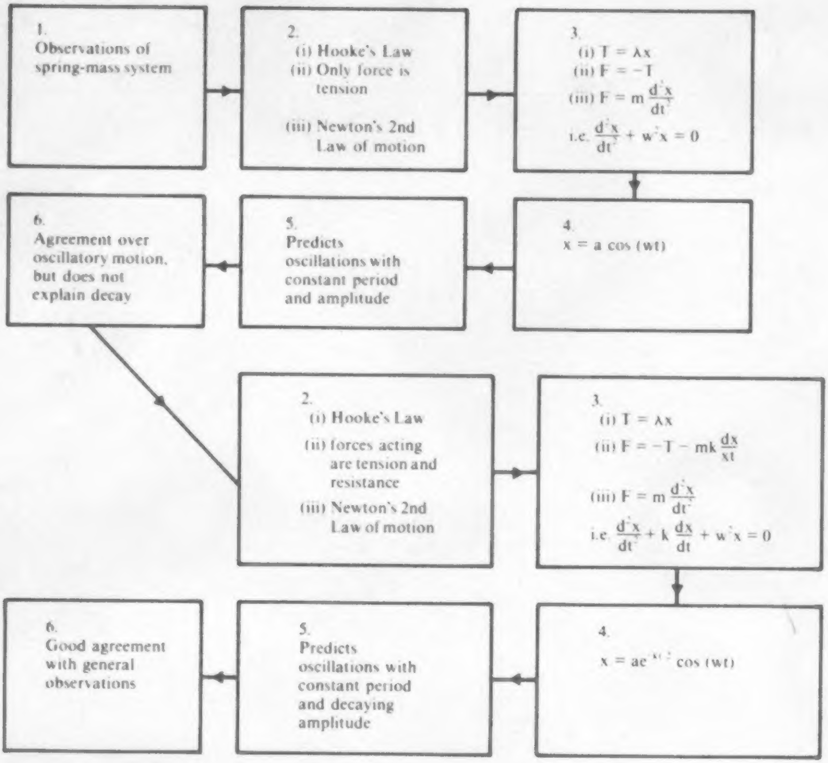

1.
Observations of spring-mass system
2.
(i) Hooke's Law
(ii) Only force is tension
(iii) Newton's 2nd Law of motion
3.
(i) $T = \lambda x$
(ii) $F = -T$
(iii) $F = m\frac{d^2x}{dt^2}$
i.e. $\frac{d^2x}{dt^2} + w^2x = 0$
4.
$x = a\cos(wt)$
5.
Predicts oscillations with constant period and amplitude
6.
Agreement over oscillatory motion, but does not explain decay
2.
(i) Hooke's Law
(ii) forces acting are tension and resistance
(iii) Newton's 2nd Law of motion
3.
(i) $T = \lambda x$
(ii) $F = -T - mk\frac{dx}{xt}$
(iii) $F = m\frac{d^2x}{dt^2}$
i.e. $\frac{d^2x}{dt^2} + k\frac{dx}{dt} + w^2x = 0$
4.
$x = ae^{-kt/2}\cos(wt)$
5.
Predicts oscillations with constant period and decaying amplitude
6.
Good agreement with general observations

Fig. 5.

In the case studies which follow we suggest that you use the boxes to illustrate the modelling process each time. You will become familiar with the concepts involved, and it will help you to gain confidence in tackling new modelling problems.

2 – King Arthur's Round Table

The famous round table fixed* in the wall in the great hall of Winchester Castle is 18ft. in diameter and is divided into 25 sectors, one for the King and then one for each of the knights, the shape being round so that no knight could claim precedence over others. It has been widely thought that this table might well be the actual celebrated table of King Arthur. In order to test this theory, the table was extracted from the wall and lowered to the ground. Scientists then got to work using a number of different techniques in order to estimate the date of the origin of the table. One of the methods employed, which was to prove all-important, was 'Carbon Dating', a technique developed by an American chemist, W. F. Libby, in the late 1940's and for which he received the 1960 Nobel Prize in Chemistry.

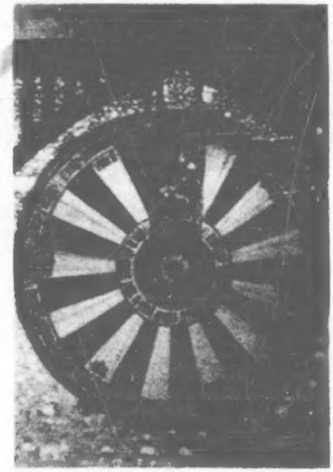


Fig. 6.

Carbon dating is based on radioactivity, a phenomenon discovered at the beginning of the century by Rutherford. Certain atoms are inherently unstable so that after some time and without external influence, they undergo a transition to an atom of a new element, with the emission of some sort of radiation. We will first show the way in which radioactivity can be modelled mathematically and then develop its use in carbon dating.

* Reproduced below by kind permission of Hampshire County Council.

Let $N = N(t)$ denote the number of atoms in a radioactive sample at time t. The derivative $\frac{dN}{dt}$ will then represent the rate of change of the number of atoms and so is the negative of the rate of disintegrations per unit time. From experimental evidence, Rutherford established that the disintegrations per unit time were approximately proportional to the number of atoms present; in mathematical terms we can express Rutherford's model as

$$\boxed{\frac{dN}{dt} = -\lambda N} \tag{1}$$

where the positive constant λ is known as the decay constant. The value of λ will be different for different radioactive elements (zero, in fact, for non-radioactive elements). The larger λ, the quicker the sample decays. λ has to be found experimentally, and in practice we measure the *half-life* of radioactive substances. This is defined as the time required for half of a given quantity of atoms to decay. Thus if at $t = 0$, $N = N_0$, we can solve to get

$$N = N_0 e^{-\lambda t} \tag{2}$$

and the half-life, τ say, is found by putting $N = N_0/2$, $t = \tau$ in (2), i.e.

$$\frac{N_0}{2} = N_0 e^{-\lambda t},$$

and taking logs,

$$\boxed{\tau = \log 2/\lambda}$$

So there is a simple relationship between the decay constant λ and half-life τ.

To indicate the possible range of values for the half-life of substances we give three examples below.

Substance	Half-life
Lead 210	22 years
Carbon 14	5568 years
Uranium 238	4.5×10^9 years

Hence, using (3), the decay constant for Carbon 14 is given by

$$\lambda = \log 2/\tau = 0.6931/5568 = 1.245 \times 10^{-4} \text{ per year.}$$

Having developed our model for radioactivity, we move onto its application in carbon dating. The earth's atmosphere is being continuously bombarded by cosmic rays, which produce neutrons in the atmosphere. These neutrons combine with nitrogen to form carbon 14, i.e. C^{14}, and this substance decays radioactively. In living tissue, the rate of absorption of C^{14} atoms is balanced by their decay and an equilibrium state is reached, giving a steady state uniform distribution of C^{14} atoms in carbon-bearing material on the surface of the earth. If a sample from this environment is isolated, the number of C^{14} atoms present in the sample begins to decay. It is assumed that the rate of bombardment of the earth's surface by cosmic rays has remained constant throughout time, so that the original rate of disintegrations of C^{14} in a charcoal sample is the same as the rate measured today.

Suppose a sample was formed at time $t = 0$. From (1) the original rate of disintegrations, $R(0)$ say, is given by

$$R(0) = \lambda N(0)$$

The present rate $R(t)$ say, from (1) and (2) is given by

$$R(t) = \lambda N(t) = \lambda N_0 e^{-\lambda t},$$

so that

$$\frac{R(t)}{R(0)} = e^{-\lambda t}.$$

Thus the age of the sample is determined from

$$\boxed{t = \frac{1}{\lambda} \log \frac{R(0)}{R(t)}} \qquad (4)$$

$R(t)$, the present rate of disintegrations of C^{14} in the charcoal sample, can be measured; and $R(0)$, the initial rate, must equal the rate of disintegrations in a comparable amount of living wood. In this way the age of the sample can be determined from (4).

We can now return to the great round table in Winchester Castle. The present rate of disintegrations from a sample was measured as

$$R(t) = 6.08$$

whilst the present rate for living wood is

$$R(0) = 6.68 \text{ (per minute per gram)}.$$

Hence the age of sample is given by

$$t = \frac{1}{1.245 \times 10^{-4}} \log \frac{6.68}{6.08} \simeq 700 \text{ years}.$$

This means that the table was built in about 1275, and so clearly could not be King Arthur's round table. There were also many other pieces of evidence that supported this conclusion so that the round table in Winchester Castle can alas be no longer called King Arthur's.

Related Problem

Charcoal from the famous Lascaux Cave in France gave an average count in 1950 of 0.97 disintegrations (per minute per gram). Living wood gave 6.68 disintegrations. Estimate the date of formation of the charcoal which therefore gives a date to the remarkable paintings in the cave.

3 – Traffic Flow

Traffic flow models can help in avoiding hold ups and maximising the flow in congested conditions. An increase in the speed of a queue of vehicles usually results in an increase in the distance between the vehicles so that greater speed does not necessarily mean greater flow of traffic. The distance a driver maintains between the front of his car and the back of the vehicle in front depends on what the driver regards as being safe. All motorists will know that there is wide variation between individual drivers!

The Highway Code lists the stopping distances for various speeds. The illustration on the back of the pamphlet is shown in figure 7. The stopping distances are made up of two parts, the thinking distance and the braking distance; we first show how these two quantities are modelled.

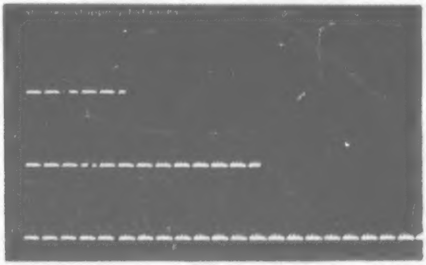


Fig. 7.

(a) Thinking distance

This is the distance the car travels *after* the driver has taken the decision to stop and before his or her foot actually operates the brake pedal. The estimated time is a little less than one second. For simplicity (as we will see later) the Ministry of Transport takes the value $\frac{15}{22}$ secs. Now if thinking time is t_t and the speed is u, then the car travels a distance

$$\boxed{d_t = ut_t} \qquad (1)$$

before the driver's foot hits the pedal.

If the speed is measured in m.p.h., then u m.p.h. = u × $\frac{22}{15}$ feet per second and

$$d_t = u \times \frac{22}{15} \times \frac{15}{22} \text{ feet} = u \text{ feet}. \qquad (2)$$

Thus a speed of 30 m.p.h. produces a thinking distance of 30 feet, etc. Note that a speed of 100 m.p.h. gives a thinking distance of 100 feet, which is quite a long way to travel before the brakes are even applied. Of course these "reaction" times are only average times. A driver who is constantly in the habit of braking and accelerating (e.g. rally drivers) might well have shorter thinking distances, whereas an elderly motorist, or a driver preoccupied with other worries, could have much longer thinking distances.

(b) **Braking Distance**

To determine the braking distance, we first suppose that the maximum braking force, say f, for any vehicle is the same at all speeds. In reality the braking force will be dependent on such factors as the road surface, tyre condition and the weather, but for general conclusions it is probably a reasonable approximation to take f as constant. The acceleration produced is given by Newton's second law of motion as

$$ma = -f.$$

Clearly different cars produce different braking forces but as a very rough approximation we suppose that the maximum braking force is 2/3 the weight of the vehicle, i.e.

$$f = \frac{2}{3} mg$$

where g is the acceleration due to gravity. Hence

$$a = -\frac{2}{3} g. \qquad (3)$$

Now if x denotes the distance travelled

$$\frac{d^2x}{dt^2} = -\frac{2}{3} g$$

and integrating

$$\frac{dx}{dt} = -\frac{2}{3} gt + A.$$

The arbitrary constant A is determined from $\frac{dx}{dt} = u$ at $t = 0$, i.e. $A = u$. Thus

$$\frac{dx}{dt} = -\frac{2}{3} gt + u, \qquad (4)$$

and integrating again with $x = 0$ at $t = 0$ gives

$$x = \frac{1}{3} gt^2 + ut. \qquad (5)$$

The time to reach zero speed is given by (4) with the left hand side zero, i.e.

$$t_b = \frac{3u}{2g},$$

and the braking distance is determined from (5) with $t = t_b$, giving

$$\boxed{d_b = \frac{3}{4}\frac{u^2}{g}.} \qquad (6)$$

Expressing u in m.p.h. we have u m.p.h. $= u \times \frac{22}{15}$ feet per second and $g = 32.2$ feet per sec. per sec. so that

$$d_b = \frac{9}{4}\left(\frac{22}{15}\right)^2 \frac{u^2}{(32.2)} = \frac{u^2}{19.96} \text{ feet}$$

so to a good approximation

$$\boxed{d_b = \frac{u^2}{20} \text{ feet}} \qquad (7)$$

and we obtain the results shown in figure 7.

The stopping distance is the thinking distance plus the braking distance; i.e. if u is the speed in m.p.h.

$$\boxed{d_s = \left(u + \frac{u^2}{20}\right) \text{ feet}} \qquad (8)$$

The Ministry of Transport has of course made many simplifying assumptions that do not apply to all vehicles and all drivers. Nevertheless they hope it will serve the purpose of giving drivers some idea of the minimum stopping distances for their cars.

We can move on to consider models of traffic flow. Let us suppose we have a constant flow of traffic in one direction in a single lane. What is the separation between the vehicles? It should really be the stopping distance d_s, but in practice many drivers are apt to use the thinking distance, d_t. The reasoning behind this is that after a car's brake lights go on it will not come to rest until it has travelled the braking distance, d_b. If the car immediately following sees the brake lights, it will stop within the stopping distance, $d_s = d_t + d_b$. So if vehicles are a thinking distance apart, they should just be able to pull up without a collision. In practice, a strategy somewhere between the two extremes is usually adopted. We will analyse the situation in the two extreme cases.

(i) MODEL 1 – SEPARATION IS THE THINKING DISTANCE

Suppose the stream of traffic is moving with speed u, and let the separation between the cars be d and the average car length be b. Then if one car has just passed a fixed point, it will take a time $(d + b)/u$ for the next car to pass the same point. So the cars pass at intervals of time $(d + b)/u$. Hence the flow rate is given by

$$f = \frac{u}{(d + b)}.$$

Now according to this model $d = u$ feet, from (1) where u is the speed in m.p.h., so that

$$f = \frac{u}{(b + u)} \times \frac{22}{15}\, s^{-1} \qquad (9)$$

Using $b = 12$ feet, the flow rate is illustrated with varying u in figure 8. There is a considerable increase in traffic flow as the speed increases, but any increase in speed above 30 m.p.h. produces little increase in flow rate.

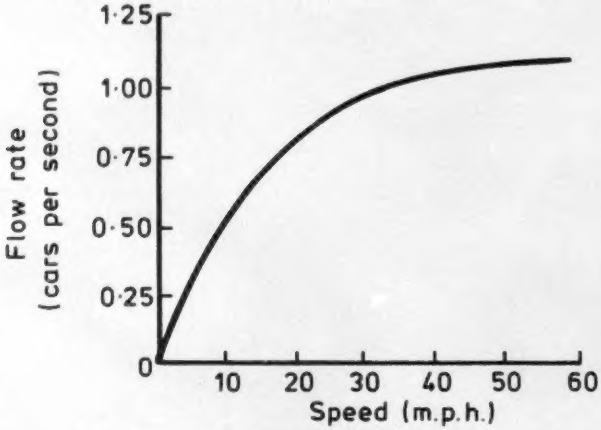


Fig. 8.

(ii) MODEL 2 – SEPARATION IS THE STOPPING DISTANCE

For this model, the separation distance d is given by (8) so that the flow rate is determined by

$$f = \frac{u(22/15)}{b + u + u^2/20} \qquad (10)$$

where b is in feet, u in m.p.h. and f is cars per second. As expected the flow rate increases, although eventually a maximum is reached. After this the flow rate slowly decreases. This is illustrated in figure 9.

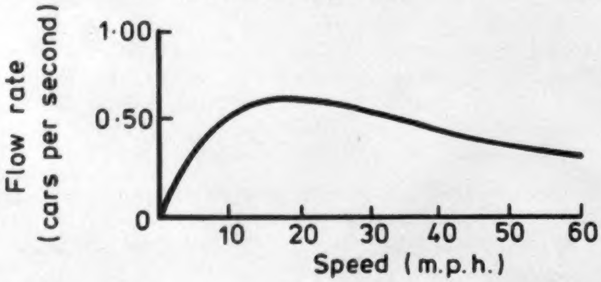


Fig. 9.

The important result here is the maximum value. We can determine where this occurs by evaluating $\frac{df}{du} = 0$. Now

$$\frac{df}{du} = \left[\frac{1(b + u + u^2/20) - u(1 + u/10)}{(b + u + u^2/20)}\right]\frac{22}{15}$$

$$= 0$$

when $b + u + u^2/20 = u + u^2/10$. This gives $u^2 = 20b$, and using $b = 12$ feet, the value of the speed which gives a maximum flow rate is $u \doteq 15$ m.p.h. This value agrees well with reality. Observations have shown that in congested conditions, maximum flow is in fact obtained with a speed of about 15 m.p.h.

A practical use of the model would be to suggest optimum speeds and separation distances for traffic through a one lane tunnel (e.g. Tyne Tunnel). In congested conditions, in order to maximise the flow rate, we would suggest

speed: 15 m.p.h.

separation distance: 26 feet.

Related Problems

1) Determine optimum speed and separation distance, when

$$d_s = d_t + \frac{1}{2} d_b$$

2) In Model 2, explain why the flow rate tends to zero as the speed increases.

3) Suggest why motorway pileups often happen under Model 1 but not under Model 2.

4 – Clunk-Click

At the beginning of this century, all motor cars had to be preceded by a pedestrian carrying a red flag. Accidents were few but travelling by car was no quicker than walking! Now there are vastly more cars on the roads and they travel at appreciably faster speeds. Consequently there are more accidents, which we cannot hope to eliminate altogether. We can, however, cut down the number of accidents by

(i) reducing the likelihood of human and mechanical faults,
(ii) reduce the number of potential hazards on the roads,
(iii) reduce the severity of the accident by better protection of the people involved.

It is the last of these three factors which we will concentrate on here.

Nearly 50% of all serious accidents involve at least one vehicle in a frontal collision. Such collisions may be head on or at an angle; the vehicle may hit another vehicle, bridge, wall, etc. One method to lessen injury in such impacts is to pad the interior and design the front section to cushion the collision. Another way is to ensure that passengers are wearing effective seat belts.

The key factor affecting the safety of the passengers is the force they experience in the impact, and whether those in the front hit the windscreen. In order to assess these problems we will consider a car in a frontal impact with a solid wall, as shown in figure 10.

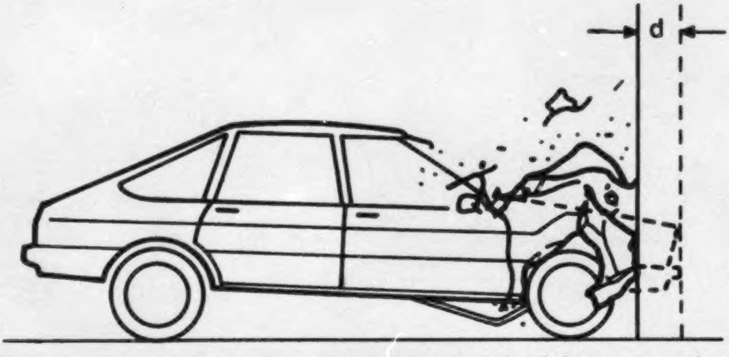


Fig. 10.

If we suppose that the car has initial speed u and that it is compressed a distance d, then we can use the equation "$v^2 = u^2 + 2as$" to find the deceleration during the impact. This is assuming that the deceleration is *constant*. Figure 11 shows some experimental results for the deceleration of a car. Although clearly not a constant, we use this approximation as we are trying to find not precise answers but general conclusions.

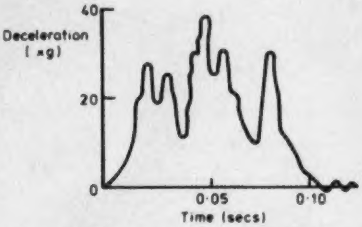


Fig. 11.

Thus with constant acceleration a, v = 0 and s = d, we have

$$a = -\frac{u^2}{2d}. \qquad (1)$$

So there is an average deceleration of $u^2/2d$, and from Newton's second law of motion, the force producing this is

$$\boxed{f = m_c u^2/2d} \qquad (2)$$

where m_c is the mass of the car. This equation does enable us to make some general conclusions. For example the force decreases as d increases. The more the front of the car is compressed, the smaller the force experienced by the passenger compartment. If d is doubled, f is halved. It is also of interest to note that if the initial speed is doubled, the force is increased by a factor 4!

We now consider what actually happens to an occupant at the front. First we consider the motion of the windscreen. This will have an acceleration $(-u^2/2d)$, and using the formula "$s = ut + \frac{1}{2}at^2$",

$$x_w = ut - \frac{u^2t^2}{4d} \qquad (3)$$

Suppose now that the passenger is *not* wearing a seat belt. Then because there is no significant force acting on the passenger after the collision, the passenger continues to move with the initial speed u of the car, until of course the windscreen is met. So if x_p measures the distances from the initial position of the windscreen, which is initially $-w$, say, then

$$x_p + w = ut. \qquad (4)$$

The variation of x_w and x_p with time are shown in figure 12. The passenger hits the windscreen when $x_w = x_p$, i.e. when

$$w = u^2t^2/4d;$$

which gives $t = (4dw)^{\frac{1}{2}}/u$. This is the time which elapses before the passenger hits the windscreen. The vital factor now is the difference between the passenger's speed and the windscreen's speed. The larger the difference, the greater the injury and the more likely that the windscreen will break.

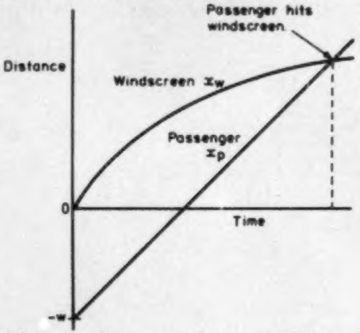


Fig. 12.

The passenger's speed is u, and we can determine the windscreen's speed from "v = u + at", i.e.

$$v = u - \frac{u^2}{2d} \cdot \frac{(4dw)^{\frac{1}{2}}}{u}$$

$$= u - u(w/d)^{\frac{1}{2}}.$$

Hence the difference in speed is

$$u(w/d)^{\frac{1}{2}}; \qquad (5)$$

so this is the impact speed. As expected an increase in d causes a decrease in the impact speed; not quite so obvious is the dependence on w, the initial distance between the passenger and the windscreen. The result shows that if a passenger is not wearing a seat belt, then it is safer to drive as near to the windscreen as possible.

Suppose now that the passenger is wearing a seat belt. This imposes a force on the passenger which will change the motion to a parabola. So long as the two parabolas describing the motion of the passenger and the windscreen do not intersect, the passenger will not hit the windscreen. This depends on the magnitude of the force exerted by the seat belt. The more the belt extends on impact, the less the force on the passenger, which means that the distance between the passenger and the windscreen should be as large as possible – the opposite to a design without seat belts.

So the moral appears to be either to drive with your nose glued to the windscreen or "clunk-click".

Related Problems

1) When does the windscreen come to rest? What is the condition for the passenger to hit the windscreen before it comes to rest? Is this likely to be realised in practise? What happens to passengers in the back seat?
2) Describe, graphically, what happens if the passenger is wearing a seat belt.

5 – Optimum Order Quantity

Consider the problem of a manufacturer who requires a definite known quantity of various components throughout, say, a year. He uses them up at a constant rate. In what quantities should he order his components? Should he make one order per year, in which case he will need sufficient storage space to stock a year's supply of components, or should he place a weekly order for 1/52 of the total requirements, in which case there will be far more paperwork and transportation costs will probably increase?

We will be more specific by looking at a car manufacturer's component problem. He needs a weekly quantity of 100 headlamps which means an annual supply of 5,000, assuming a two-week closure of the factory in the year. The administration and transportation costs incurred by placing one order vary very little with the actual size of the order and are estimated at about £60 per order, whereas the storage costs and loss of revenue due to capital being tied up is calculated at £2 per headlamp per year. The manufacturer requires to know the order quantity which will minimise his total operating costs per year.

One of the advantages of a mathematical model is that it is often possible to formulate it in general terms, solve it, and at the last stage refer back to a specific problem. It can then be used again and again for other problems of the same type. We illustrate this technique here by formulating a general model for the operating costs that depends on three factors, namely

(i) cost of the components,
(ii) cost of placing orders,
(iii) cost of carrying stock.

Suppose the yearly requirement is A units, each costing p pounds. The total yearly cost of this component is then given by

$$C_c = Ap. \tag{1}$$

Let us suppose that there are n orders per year each of size $x = A/n$. If t is the cost of placing one order, no matter the size of the order, the total yearly ordering costs are given by

$$C_o = nt = At/x. \tag{2}$$

At the beginning of the year, the manufacturer starts with a stock quantity x, which decreases steadily to zero, before returning to level x when the next order is received. This pattern continues throughout the year, and is illustrated in figure 13 below.

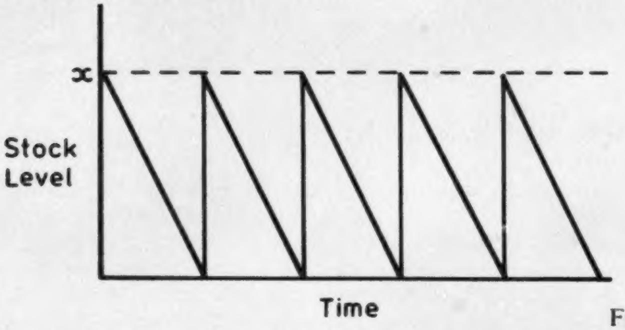


Fig. 13.

The average stock held throughout the year is $x/2$, so that if s is the stocking cost per item per year, the total yearly stocking costs are given by

$$C_s = sx/2. \tag{3}$$

The total yearly operating costs are given by the sum of the costs for components, ordering and stocking; i.e.

$$C = C_c + C_o + C_s = Ap + At/x + sx/2 \tag{4}$$

The manufacturer requires the batch size x which minimises the total costs C. Using calculus, we need to solve

$$\frac{dC}{dx} = 0$$

which gives

$$-\frac{At}{x^2} + \frac{s}{2} = 0;$$

so that

$$\boxed{x = \sqrt{2At/s}} \tag{5}$$

This quantity is known as the economic order quantity (E.O.Q.) since it is the order size which minimises the total costs.

The general problem has been solved, so we can return to the specific one introduced earlier. The value of the parameters used in the model are

$$A = 5000,\ t = 60,\ s = 2$$

so that

$$\text{E.O.Q.} = \sqrt{\frac{2\,.\,5000\,.\,60}{2}} \simeq 547.$$

The number of orders placed per year, n, is given by

$$n = \frac{A}{x} = \frac{5000}{547} = 9.14.$$

In practice this must of course be a whole number, so that the optimum number of orders will be 9 or 10. The manufacturer might well be likely to choose 10, of order size 500, since this is a very convenient order size. The answer 9.14 is not to be taken as a precise answer but just as an estimate of the likely optimum order size. The assumptions in the model are only approximations to the real problem; ordering costs and stocking costs are just estimates. So the model should only be used to give an indication of the answer, and the car manufacturer must then use his judgement and intuition in the choice of the order size. He might well, for instance, never like to see his stock level fall below 200 in case of strikes in the component manufacturers.

Related Problems

1) Extend the model to allow for running out of stock.

2) What happens if discounts are offered? For example, suppose the basic price is £1 per headlight, but the following discount is offered
£0.90 per headlight for orders of 1000 or more
£0.80 per headlight for orders of 2000 or more.
What is the E.O.Q.?

6 – Single Species Population Models

The problem of estimating population changes is clearly an important one. The fact that the birth rate in the U.K. has been in a decline for over a decade has consequences over the next few decades in terms of the country's resources. Fewer maternity beds are required at the first stage, less medical facilities for young children are needed, and the number of new children entering primary school each year decreases. If classes are kept the same size, fewer teachers are required and consequently fewer need to be trained, which means less teacher training facilities and so on. But such changes in the way the country's resources are spent need to be planned well ahead, and so good estimates of how the population is changing, not just the overall size, but the age distribution as well, are needed. Similarly towns and cities need to estimate population changes so that suitable facilities and amenities are available.

Mathematical models used to answer these sorts of questions are complex, so in this article two models will be described which attempt to describe the way in which the total population of a country changes. The first one is called the Malthusian population model, since it is based on ideas put forward by the English economist Thomas Malthus in his article "An Essay on the Principle of Population" published in 1798. If $N = N(t)$ denotes the country's total population at time t, then in a small time interval, say δt, it is assumed that both births and deaths are proportional to the population size and the time interval, i.e.

$$\text{births} = \alpha N\delta t, \quad \text{deaths} = \beta N\delta t \ (\alpha, \beta \text{ constants}).$$

Thus the increase, say δN, in the total population in the time interval δt is given by

$$\delta N = \alpha N\delta t - \beta N\delta t = (\alpha - \beta)N\delta t = \gamma N\delta t$$

where $\gamma = \alpha - \beta$. Dividing by δt and taking the limit as $\delta t \to 0$ leads to the differential equation

$$\boxed{\frac{dN}{dt} = \gamma N.} \tag{1}$$

This has solution

$$N = N_0 e^{\gamma t} \tag{2}$$

where at $t = 0$, $N = N_0$. The predicted behaviour of the population depends very much on the sign of the constant γ. We have exponential growth if $\gamma > 0$, decay if $\gamma < 0$ and stationary if $\gamma = 0$. These solutions are illustrated in figure 14.

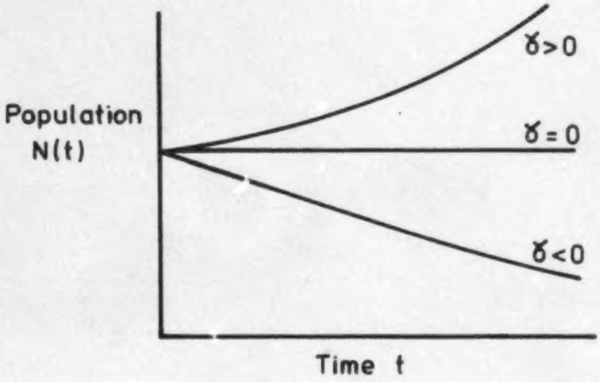


Fig. 14.

In order to see if this model is of any value in population predictions, we turn to a specific problem. In table 1 the U.S.A. population statistics are given for 1790, 1800 and 1810. We will use the Malthusian model to predict the future U.S.A. population total. Time $t = 0$ corresponds to 1790, so we take $N_0 = 3.9 \times 10^6$. Working in time intervals of 10 years, at $t = 1$ (i.e. 1800), from (2),

$$N(1) = 5.3 \times 10^6 = 3.9 \times 10^6 \, e^{\gamma} ,$$

so that

$$\gamma = \log (5.3/3.9) = 0.307 .$$

Year	U.S.A. Population ($\times 10^6$)
1790	3.9
1800	5.3
1810	7.2

Table 1

Using these values for N_0 and γ, the model predicts $N(2) = 7.3 \times 10^6$ which is reasonably close to the observed value at 1810. Continuing in this way we calculate the predicted population total every 10 years. These values, together with the actual values, are given in table 2.

Year	U.S.A. Population ($\times 10^6$)	Predicted Values ($\times 10^6$)
1820	9.6	10.0
1830	12.9	13.7
1840	17.1	18.7
1850	23.2	25.6
1860	31.4	35.0
1870	38.6	47.8
1880	50.2	65.5
1890	62.9	89.6
1900	76.0	122.5
1910	92.0	167.6
1920	106.5	229.3
1930	123.2	313.7

Table 2

There is reasonable agreement for some time. For example the 1850 figure has a 10% error. But by 1870 the error has increased to 30% and the model is of little use.

Having been through the modelling cycle once, if we are not satisfied with the model, we must look back at the assumptions made and consider what factors have been neglected. As it stands the Malthusian model predicts unlimited growth (if $\gamma > 0$) for all future time. This is most unlikely ever to occur since there are many varied limitations to growth, such as lack of food resources, overcrowding, insufficient energy supply, and pollution destroying the production environment. Is it possible to represent these factors in an equation? In 1837 Verhulst proposed an extension to the Malthusian model which took into account "crowding" factors with an extra term on the right hand side of the governing differential equation. His model is given by

$$\boxed{\frac{dN}{dt} = \gamma N - \lambda N^2} \tag{3}$$

where λ is a positive constant. For small N, the dominant term on the right side is γN and we have the original Malthusian model, but as N increases, the $-\lambda N^2$ term becomes more important and so restricts the growth of the population.

Verhulst's model is a first order differential equation which can be solved by separating variables, i.e.

$$\int \frac{dN}{N(1 - \lambda N/\gamma)} = \int \gamma \, dt$$

$$\int \left[\frac{1}{N} + \frac{(\lambda/\gamma)}{(1 - \lambda N/\gamma)}\right] dN = \gamma t + K$$

where K is a constant. Thus

$$\log N - \log(1 - \lambda N/\gamma) = \gamma t + K$$

and if, at $t = 0$, $N = N_0$, we have

$$K = \log [N_0/(1 - \lambda N_0/\gamma)].$$

Hence

$$\frac{N}{(1 - \lambda N/\gamma)} = \frac{N_0 e^{\gamma t}}{(1 - \lambda N_0/\gamma)}$$

and solving for N gives

$$N(t) = N_0 \gamma e^{\gamma t}/[\gamma + \lambda N_0(e^{\gamma t} - 1)]. \tag{4}$$

What does the model predict? The major change from the Malthusian model is that there is not unlimited growth. In fact as $t \to \infty$,

$$N \to \gamma/\lambda .$$

A sketch of a typical solution is given in figure 15.

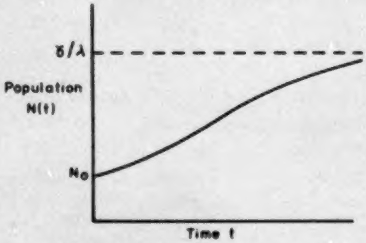


Fig. 15.

We now move onto the validation stage of the modelling process. Again using the U.S.A. population data given in Table 1, Verhaulst chose the parameters as

$$N_0 = 3.9 \times 10^6, \quad \gamma = 0.3134, \quad \lambda = 1.5337 \times 10^{-9}.$$

The predicted values are given in Table 3 together with the actual values up to 1930. There certainly appears to be a remarkable correlation, since for over 100 years the Verhaulst model accurately predicted the U.S.A. population.

Year	U.S.A. Population ($\times 10^6$)	Predicted Values ($\times 10^6$)
1820	9.6	9.7
1830	12.9	13.0
1840	17.1	17.4
1850	23.2	23.0
1860	31.4	30.2
1870	38.6	38.1
1880	50.2	49.9
1890	62.9	62.4
1900	76.0	76.5
1910	92.0	91.6
1920	106.5	107.0
1930	123.2	122.0

Table 3

So as far as this situation goes, i.e. the prediction of the total U.S.A. population after 1810, the model looks an excellent one. It does though go astray after 1930. The limiting value of the Verhaulst model is 197×10^6, but already the U.S.A. population is past 200×10^6. Of course, it is perhaps a little unreasonable to expect the model to predict accurately for more than 100 years. On the other hand, can we have much confidence in using this model today to predict the future changes for a country? The laws governing population changes are not as precise as those that govern a mechanical system, and this is why it is very difficult to have any real confidence in any population predictions for the future. We do not know what fundamental changes might take place in the next few decades – a nuclear war, a population catastrophe or another industrial revolution.

Related Problems

1) An alternative way to derive the Verhulst population model is to assume that there is some upper limit to the population N which can be sustained, say N_∞. Then the rate of increase of population is assumed to be proportional to both N (Malthusian Law) and the degree of unsaturation $(1 - N/N_\infty)$. Show that this is equivalent to equation (3).
For the American population data given above, what is the value of N_∞ implied by the model?

2) The Irish population data for 1801-1916 is given below. Develop a model to describe this situation. It should be noted that there were potato famines in the middle of the 19th Century.

Year	Population (in thousands)	Year	Population (in thousands)
1801	5216	1861	5788
1806	5574	1866	5523
1811	5956	1871	5398
1816	6365	1876	5278
1821	6802	1881	5146
1826	7269	1886	4906
1831	7767	1891	4680
1836	7981	1896	4542
1841	8200	1901	4447
1846	8288	1906	4391
1851	6514	1911	4381
1856	5973	1916	4273

7 - Launching a Satellite

The launching of yet another satellite no longer provides the news headlines (unless it is a failure); it is just a routine part of our life now, very unlike the 4th October 1957 when the first earth satellite, Sputnik 1, was launched by Russia. Since then rocket technology has progressed at an enormous pace, partly of course due to the military possibilities of space flight but partly for prestige and also due to man's pursuit of knowledge. The Saturn rockets used for the Appollo moonshots were essentially three-stage rockets, together with the satellite and its lunar module. In this article we will show why single stage rockets are of little use in launching an earth satellite into a typical orbit. We will also find how a two-stage rocket should be designed in order to maximise the final speed given to the satellite.

Initially we will determine the speed that a satellite needs in order to stay in a circular orbit of height h above the earth's surface, radius R_e. The gravitational pull towards the centre of the earth is given by Newton's inverse square of attraction as

$$\gamma m M_e / a^2$$

where γ is the universal gravitational constant, m the mass of the satellite, M_e the earth's mass and $a = h + R_e$, the radius of the orbit. This force is balanced by the centrifugal force mv^2/a, v being the satellite's speed. Thus

$$\gamma m M_e / a^2 = m v^2 / a$$

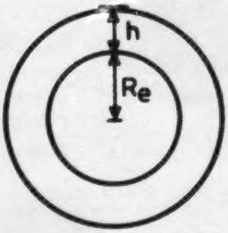
giving

$$\boxed{v = \gamma M_e/(h + R_e)\ ^{\frac{1}{2}}} \qquad (1)$$

Now for a typical orbit of height 100 km above the earth's surface,

$$V = 7.8\,\mathrm{km\ s^{-1}}. \qquad (2)$$

So in order to launch such a satellite, the rocket must be able to reach speeds of this magnitude.



Let us now consider the capabilities of single stage rockets. We start with the equation of motion, but here we must be very careful. Rockets are continuously losing mass, in fact not just losing it but it is being propelled away at significant speeds. We cannot directly apply Newton's second law (force is proportional to rate of change of linear momentum) but consider a rocket of mass m, moving with speed v which in a small time δt loses a small mass, say δm_p, which leaves the rocket with speed u in the opposite direction to v. The situation is illustrated in figure 16. The resulting speed of the rocket is $v + \delta v$.

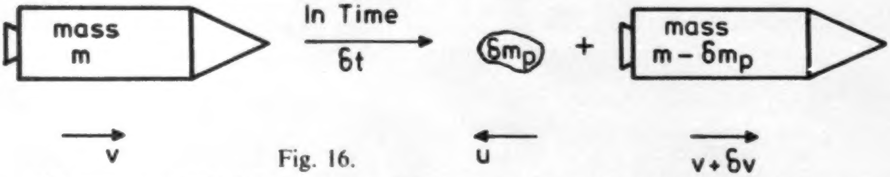


Fig. 16.

We can now apply Newton's second law of motion to the whole system, giving

$$\begin{aligned}\text{force} &= \lim_{\delta t \to 0} \left\{ \frac{\text{Change in momentum in time } \delta t}{\delta t} \right\} \\ &= \lim_{\delta t \to 0} \left\{ \frac{[(m - \delta m_p)(v + \delta v) + \delta m_p(-u)] - mv}{\delta t} \right\} \\ &= \lim_{\delta t \to 0} \left\{ m\frac{\delta v}{\delta t} - (u + v)\frac{\delta m_p}{\delta t} - \frac{\delta m_p}{\delta t}\delta v \right\}\end{aligned}$$

i.e.

$$\boxed{F = m\frac{dv}{dt} - (u + v)\frac{dm_p}{dt}}\,, \tag{3}$$

where F is the force acting on the system in the v direction. We usually write $c = u + v$, so that c is the speed of the propellant (burnt gases) relative to the rocket called the relative exhaust speed. Also the term $\frac{dm_p}{dt}$ is the positive rate of change of the propellant mass, and so

$$\frac{dm_p}{dt} = -\frac{dm}{dt}.$$

Thus for motion in one direction, (3) can be written as

$$F = m\frac{dv}{dt} + c\frac{dm}{dt}. \tag{4}$$

But what are the main forces acting on the rocket as it leaves the earth's surface? The gravitational pull is the dominant one, and in the initial motion the drag due to the earth's atmosphere will be important although this diminishes fairly rapidly with height above the surface. On the other hand we are just trying to get an overall picture of rocket flight, so initially we neglect all external force; i.e. $F = 0$ in (4). Hence our model is governed by the differential equation

$$\frac{dv}{dt} = -\frac{c}{m}\frac{dm}{dt}. \tag{5}$$

We can integrate this to get

$$\begin{aligned} v &= -c\int \frac{1}{m}\frac{dm}{dt}\,dt \\ &= -c\int \frac{1}{m}\,dm \\ &= -c\log m + k\,, \end{aligned}$$

k being the constant of integration. Initially suppose $v = 0$ and $m = m_0 + P$, where P is the satellite's (payload) mass and m_0 the initial rocket mass (excluding the satellite). This mass is divided into two parts, the initial fuel mass, say ϵm_0 $(0 < \epsilon < 1)$, and the casing and instruments of mass $(1 - \epsilon)m_0$. We call $(1 - \epsilon)$ the *structural factor* of the rocket. The constant k is determined from $0 = -c\log(m_0 + P) + k$, so that

$$v = -c\log\left[m/(m_0 + P)\right].$$

When all the fuel has been burnt, $m = (1 - \epsilon)m_0 + P$, so that

$$\boxed{v_1 = -c\log\left[1 - \epsilon/(1 + \beta)\right]}$$

where $\beta = P/m_0$.

This final speed is very much an upper bound to the possible speed that can be given to the satellite. The forces that have been neglected all reduce this final speed. Nevertheless it is instructive to evaluate (6) for a typical rocket. The predicted final speed (6) depends on the three parameters

$$c,\ \epsilon \text{ and } \beta.$$

The relative exhaust speed c, for both liquid and solid fuel rockets, has a typical value of about 3.0 kms s^{-1}. The value of ϵ depends on the materials used but is typically about 0.8. The value of β, the ratio of the payload mass to the rocket mass, will be small, say about 1/100. With these values, (6) gives

$$v_1 \simeq 4.7 \text{ km s}^{-1},$$

which is nowhere near the required value for launching a satellite. If gravity is included in the model it reduces the final speed by about 4%, and so the neglect of it in the analysis is of no real consequence.

Our conclusion is that one stage rockets are virtually incapable of launching a typical earth satellite. So we turn to multi-stage rockets in which after each stage has been burnt, the casing of that stage is discarded. In this way the unwanted casing mass is not carried throughout the motion and so we would clearly expect a better final speed. How much better though? Let's consider a two stage rocket, with stages of initial mass m_1 and m_2 respectively. For simplicity assume equal relative exhaust speeds, c, and equal structural factors.

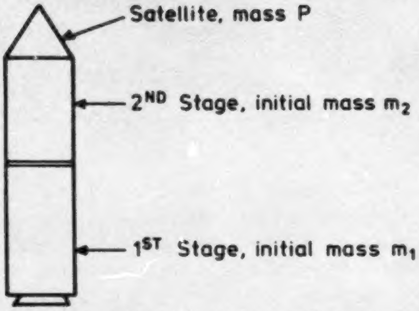


Fig. 17.

For the first stage burn the payload is the second stage and the satellite, i.e. $m_2 + P$. Thus from (6), with $\beta = (m_2 + P)/m_1$, the speed obtained by the first stage is given by

$$- c \log [1 - \epsilon m_1/(m_1 + m_2 + P)].$$

The second stage is just a single stage rocket with initial mass m_2 and payload P, and again using (6), this stage provides a further increment in speed

$$- c \log [1 - \epsilon m_2/(m_2 + P)].$$

Thus the final speed given to the satellite is given by

$$v_2 = - c \log [1 - \epsilon m_2/ m_1 + m_2 + P)] - c \log [1 - \epsilon m_2/(m_2 + P)] \quad (7)$$

To compare with the single stage rocket considered above we take, for example

$$\epsilon = 0.8,\ c = 3.0\,\mathrm{km\,s^{-1}},\ m_1 = m_2 = 50P$$

so that $P/(m_1 + m_2) = 1/100$; and with these values

$$v_2 \simeq 6.2\ \mathrm{km\ s^{-1}}.$$

So although we have obtained a considerable improvement over the single stage rocket ($v_1 \simeq 4.7\ \mathrm{km\ s^{-1}}$) it is still not sufficient to launch the satellite.

But before rushing to a 3 stage rocket, there is one more action we can try. In the values used above we took equal stage masses, i.e. $m_1 = m_2$. Is there a better choice for the ratio $m_1 : m_2$ than $1 : 1$? Will the final speed vary if we keep the total mass $m_1 + m_2 = m_0$, say, constant but allow the ratio $m_1 : m_2$ to vary? The answer is yes, and to see how to best choose the ratio $m_1 : m_2$, put $m_1 = m_0 - m_2$ in (7). Then

$$v_2 = - c \log [1 - \epsilon(m_0 - m_2)/(m_0 + P)] - c \log [1 - \epsilon m_2/(m_2 + P)]$$

and so v_2 is a function of m_2. For maximum v_2,

$$\frac{dv_2}{dm_2} = 0,$$

which gives

$$\frac{\epsilon/(m_0 + P)}{[1 - \epsilon(m_0 - m_2)/(m_0 + P)]} = \frac{\epsilon P/(m_2 + P)^2}{[1 - \epsilon m_2/(m_2 + P)]}.$$

Cross multiplying and simplifying gives

$$(1 - \epsilon)\,(m_2^2 + 2Pm_2 - Pm_0) = 0\,.$$

Since $\epsilon \neq 1$, we must have

$$m_2^2 + 2Pm_2 - Pm_0 = 0\,,$$

and solving

$$m_2 = -P + (P^2 + Pm_0)^{\frac{1}{2}}\,.$$

Thus with $\beta = P/m_0$,

$$\frac{m_2}{m_0} = -\beta + (\beta + \beta^2)^{\frac{1}{2}}$$

and since $m_1 = m_0 - m_2$,

$$\frac{m_1}{m_0} = 1 + \beta - (\beta + \beta^2)^{\frac{1}{2}} = (1 + \beta)^{\frac{1}{2}}\left[(1 + \beta)^{\frac{1}{2}} - \beta^{\frac{1}{2}}\right].$$

Thus for maximum final speed, we must choose the ratio m_1/m_2 as

$$\frac{m_1}{m_2} = \frac{(1 + \beta)^{\frac{1}{2}}}{\beta^{\frac{1}{2}}}$$

$$\text{i.e.} \quad \boxed{m_1/m_2 = \left(1 + \frac{1}{\beta}\right)^{\frac{1}{2}}} \qquad (8)$$

So with $\beta = 1/100$, the optimum ratio $m_1/m_2 = 10.05$, showing that to maximise the final speed given to the satellite, the first stage must be about ten times larger than the second. Using this new ratio, and with $\epsilon = 0.8$, $c = 3.0$ km s^{-1} as before, (7) now gives

$$v \simeq 7.65 \text{ km s}^{-1}\,.$$

At last we have reached a value close to the required one. So it appears likely that two stage rockets, provided the stage mass ratios are correctly chosen, will be able to launch satellites into earth orbit. Clearly, further staging will improve the final speed, but at the expense of further complexity and increased costs.

Related Problems

1) Introduce gravity into the model described above for a single stage rocket. What difference does it make to the estimated speed given to the payload? Is it significant?
2) Optimise the mass ratio m_1/m_2 for maximum final speed when the structural factors of the rocket stages are not necessarily equal. Check that equation (8) is still obtained when the structural factors are equal. If $\epsilon_1 = 0.9$ and $\epsilon_2 = 0.7$ what is the optimum ratio for m_1/m_2? (using the data $\beta = 1/100$).
3) Extend the analysis to three stage rockets.

8 – "Save It" – Water Heating and Cooling

In this section we are concerned with finding optimal policies for the way in which an immersion heater is operated. Is it, for example, cheaper to switch the immersion heater off at night and then on again shortly before it is needed in the morning or allow the thermostat to control the temperature all night? Before looking at the way water cools we first model the way in which water is heated.

Suppose we start with a metal tank full of cold water which is heated by an electric heater fully immersed in the water. To model this situation we will assume the temperature of the water is the same everywhere inside the cylinder. In practice the water will be at a higher temperature near the heater, which often only reaches half way down the cylinder (see figure 18).

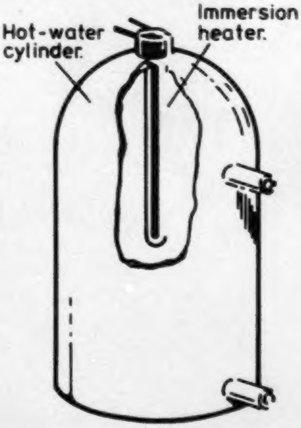


Fig. 18.

We also suppose that heat is produced at a constant rate q for a time T so that the total heat supplied is given by

$$h = qT .$$

Experiments show that the heat required to change the temperature of a mass m of liquid by θ is proportional to both m and θ

$$\text{i.e.}\quad h \propto m\theta .$$

Thus $qT \propto m\theta$ and the constant of proportionality for the liquid is called the specific heat capacity, denoted by c. Hence

$$\boxed{h = qT = cm\theta} \tag{1}$$

Now heat is a form of energy, and the S.I. unit for energy is the Joule, J. The rate at which energy is produced is called the power, and if energy is supplied at the rate of 1 Joule per second it is said to have power 1 watt, W; i.e.

$$1\ \mathrm{Js^{-1}} = 1\ \mathrm{W}. \tag{2}$$

(For example a 2 kW electric kettle produces 2000 J each second.) From (1) we see that the units for c are $\mathrm{J\ kg^{-1}\ C^{-1}}$ (where C denotes degrees Centigrade). For water

$$c = 4200\ \mathrm{J\ kg^{-1}\ C^{-1}} \tag{3}$$

and we can estimate how long it takes an immersion heater, rated at 3 kW, holding 100 kg of water to heat from 15° C to 60° C. Substituting these values in (1) gives

$$T = 4200 \times 100 \times 45/3000 \text{ sec} = 1\tfrac{3}{4} \text{ hours}.$$

Does this value check with reality? A typical cylinder holds about 100 kg of water, but the time obtained above is probably too long. So what assumptions have we made in the modelling process which were not accurate? We have taken no account of the cylinder's casing, but including it in the model will increase the time taken! Another approximation made was to assume that all the water in the cylinder is heated, but as illustrated in figure 1, often only about ½ the cylinder is heated. Since cold water tends to sink, hot to rise, the heater is placed in the *top* of the cylinder and in effect only heats say a mass m = 50 kg of water. This will give a time of heating

$$T = 52\tfrac{1}{2} \text{ mins}.$$

which is in better agreement with reality.

To bring the model even nearer to reality we must take into account the casing of the cylinder and the surroundings. Now

heat supplied by heater = heat given to water + heat given to
casing + heat given to surroundings.

If the casing is assumed to be at the temperature of the water which it is in contact with, and m_c is the mass of that part of the casing surrounding the heated water, then

$$h = c_w m_w + c_c m_c \tag{4}$$

where suffix w refers to the water and c_c is the specific heat capacity of the casing, and we have neglected heat lost to the surroundings. For example if $c_c = 400\ \mathrm{J\ kg^{-1}\ C^{-1}}$, and $m_c = 5$ kg, then the time taken to raise the water temperature from 15°C to 60°C is now given by

$$T = (4200 \times 50 + 400 \times 5) \times 45/3000 \text{ sec} = 53\tfrac{1}{2} \text{ mins}.$$

So consideration of heat lost to the casing appears to be relatively insignificant.

We are now in a position to consider the most economical way to operate an immersion heater. To analyse the problem we first consider what happens when the heater is switched off and the water cools. To model the cooling of water we suggest you try the following experiment.

Experiment

Take a cup of very hot water and place a thermometer in the liquid. Read the temperature initially and thereafter every 20 minutes,for two hours. Plot the data on a graph, drawing a smooth curve to connect the data points.

Measure the slope of the curve at every 20 minute point and plot this against the temperature.

The final graph should give an approximate straight line and illustrates Newton's law of cooling, which states that the temperature difference, θ, between the hot body and its surrounding decreases at a rate proportional to the temperature difference. In mathematical terms

$$\boxed{\frac{d\theta}{dt} = -k\theta} \qquad (5)$$

where k is a positive constant. Returning to the situation of the water cooling inside the immersion heater, we must find the appropriate form for the constant k.

If q now stands for the rate of loss of heat, an alternative way of expressing the law of cooling is to assume that q is proportional to the surface area, a, and the temperature difference θ. Thus

$$q = ja\theta$$

where the constant j is called the heat transfer coefficient, which depends on the nature of the surface of the hot body and the surrounding environment. Now if h is the total heat originally supplied, then

$$\frac{dh}{dt} = -ja\theta .$$

But from (4), $h = (c_w m_w + c_c m_c)\theta$ giving

$$\frac{d\theta}{dt} = \left[\frac{-ja}{c_w m_w + c_c m_c}\right]\theta . \qquad (6)$$

So the parameter k in (5) has value $ja/(c_w m_w + c_c m_c)$.

The solution of (6) is illustrated in figure 19. At $t = 0$, it is assumed that $\theta = \theta_0$ giving

$$\boxed{\theta(t) = \theta_0 e^{-kt}} \qquad (7)$$

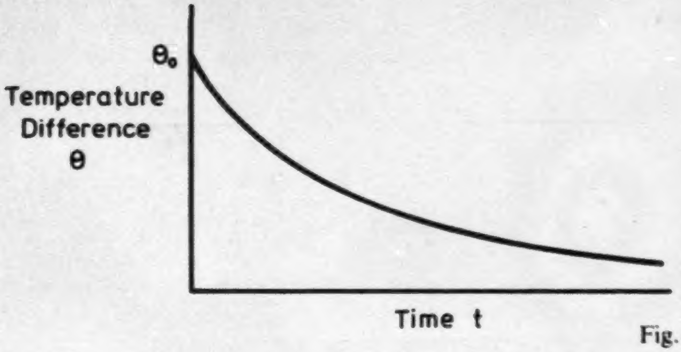


Fig. 19.

Knowing that θ decays exponentially, we can see that the slope of the curve is much greater when θ is large than when it is small. Hence keeping the water hot means that the water is always cooling at its greatest rate, the loss of heat being rectified by regular boosts from the heater. On the other hand, allowing the water to cool until just before hot water is needed again means that the cooling takes place progressively more slowly, shown by the slope of the curve becoming less steep. The loss of heat over the period of time must therefore be less, and so less heat in total needs to be supplied by the immersion heater.

So by graphical arguments it is clear that it is cheaper to switch off the heater until the water is needed. But how much cheaper?

Method 1: Suppose the immersion heater as described earlier has been maintaining the water in the cylinder at 45°C above the surrounding temperature. If it is switched off at 11 p.m. at night, then from (7), the temperature difference at 7.00 a.m. in the morning is given by

$$45e^{-28800k}$$

and if $a = 1\ \mathrm{m}^2$, $j = 10\ \mathrm{W\ m^{-1}\ C^{-1}}$, then using values as above

$$k = 10 \times 1/(50 \times 4200 + 5 \times 400) = \frac{1}{21200}.$$

Hence the temperature difference is approximately 12°C, showing that the water has dropped by 45 − 12 = 33°C.

Using (4) and (1), the time taken to raise the temperature by 33°C is

$$T = \frac{(c_w m_w + c_c m_c)\theta}{q} = \frac{(4200 \times 50 + 400 \times 5)33}{3000}\ \mathrm{s}$$

which gives $T \simeq 40$ min. A 3 kW heater uses $3 \times \frac{2}{3} = 2$ units of electricity when switched on for 40 minutes (*one unit of electricity* is consumed by a device rated at 1 kW in one hour). So this method uses approximately *2 units* of electricity.

Method 2: This time we allow the thermostat to control the temperature throughout the night. Suppose that the thermostat switches on at 42°C above the environment and off again at 45°C. The time taken to raise the temperature by 3°C is given by

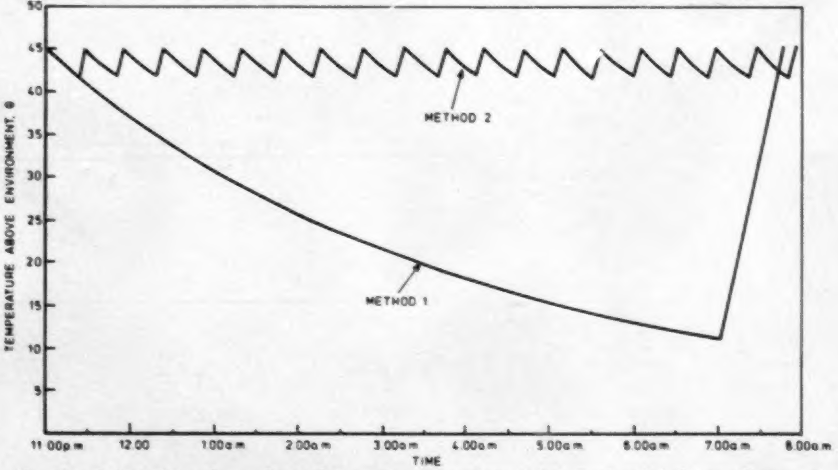
$$T = \frac{(c_w m_w + c_c m_c)\theta}{q} = \frac{(4200 \times 50 + 400 \times 5)\,3}{3000}\ \mathrm{s}$$

i.e. $T \simeq 3\frac{1}{2}$ mins. The time taken to cool from 45°C to 42°C is given from (7) by

$$t = \frac{1}{k} \log \frac{\theta_0}{\theta(t)} = 21200 \log \left(\frac{45}{42}\right)$$

i.e. $t = 24\frac{1}{2}$ mins. So the system is continually going through a cooling-heating cycle every 28 mins. This is illustrated in figure 20.

Fig. 20.



In the 8 hrs 40 mins when the system regulated by method 1 is cooling and reheating once, the system regulated by method 2 has gone through about 19 cycles, and so the heater has been on for $19 \times 3\frac{1}{2}$ minutes $= 66\frac{1}{2}$ minutes. Thus the 3 kW heater uses $3 \times 66.5/60$ units $= 3.325$ units of electricity.

So the first method saves at least one unit of electricity each night, which is about 5p a unit. Perhaps not a very significant saving, but worth £18 a year!

Related Problem

Estimation of Time of Murder

The body of a murder victim was discovered at 11.00 p.m. one evening. The police doctor on call arrived at 11.30 p.m. and immediately took the temperature of the body, which was 94.6°F. He again took the temperature after one hour when it was 93.4°F, and he noted that the temperature of the room was a constant 70°F. Use the law of cooling (as described in this Section) to estimate the time of death.

– Conclusions

In the case studies above we have seen how mathematics can be a vital tool in solving problems. Mathematical models for a variety of situations have been described. In some cases, for example the dating of the round table at Winchester, the modelling process produced definite answers. In others, for example the optimum order quantity, the mathematics produces a definite answer but it is doubtful if this particular answer has great relevance to reality. The model used has too many simplifying assumptions, and must be regarded as only a first step in the modelling cycle. We have seen, as in the rocket staging problem, how a mathematical model can be used in order to achieve an optimum *design*, whilst in the population modelling, the models are used to *predict* future values.

In all cases the important point to note is that in applying mathematics, the mathematical answers obtained need interpretation. Experience and expertise is needed to formulate a mathematical model, mathematical skill is required to solve the mathematical problem and critical judgement must be employed in deciding whether a model is capable of meeting the required degree of accuracy. The whole modelling process requires many talents, and mathematicians capable of performing such tasks will be invaluable to the future of mankind.

Opportunities for cooperation through 'modelling' are numerous. Several of the case studies in this booklet arise directly from the 'New Geography' and from Biology. All the sciences, however, provide numerous examples of 'physical' or 'real' problems for which a modelling approach is ideal. Why not talk about this with your science (or mathematics) colleagues? There is a great potential for broadening the scope of mathematics with many genuine science problems, and for applying mathematics more in the sciences. This potential can only be realised if individual teachers in schools get together to discuss this cooperation and actually cooperate. It's up to you!

Finally, for those about to embark on a modelling exercise in the classroom, good luck. Do not despair and try to keep calm, even if events are not going in the direction you planned. It is not easy to teach mathematical modelling and you will find that it will take time and experience to teach material of this sort effectively. Nevertheless we feel it is worth persevering. The possible achievements are numerous. Your pupils will see the relevance of mathematical topics, they will learn to think constructively, they will gain insight into critical judgement and, given time, patience and understanding, they will gain confidence in their own ability to solve problems.

10 - References

The material described in this booklet has been taken from a number of sources. We list below books both on mathematical modelling and also on applications of mathematics to particular topics.

1. ANDREWS, J. G. and McLONE, R. R., 1976, Mathematical Modelling (Butterworth).
2. BAILEY, N. T. J., 1975, The Mathematical Theory of Infectious Diseases (Griffin).
3. BATTERSBY, A., 1966, Mathematics in Management (Penguin).
4. BENDER, E. A., 1978, An Introduction to Mathematical Modelling (Wiley).
5. BITTINGER, M. L., 1976, Calculus, A Modelling Approach (Addison-Wesley).
6. BRADFORD, M. G. and KENT, W. A., 1977, Human Geography (Oxford).
7. BRAUN, M., 1975, Differential Equations and their Applications (Springer Verlag).
8. BURGHES, D. N. and DOWNS, A. M., 1975, Modern Introduction to Classical Mechanics and Control (Ellis Horwood).
9. BURLEY, D., 1975, Mathematical Model for a Kidney Machine, Mathematical Spectrum **8**, 69-75.
10. DERRICK, W. R., and CROSSMAN, C. J., 1977, Elementary Differential Equations with Applications (Addison Wesley).
11. GANDOLFO, G., 1971, Mathematical Methods and Models in Economics and Dynamics (North Holland).
12. GROSSMAN, S. and TURNER, J., 1974, Mathematics for the Biological Sciences (Macmillan).
13. HABERMAN, R., 1977, Mathematical Models (Prentice Hall).
14. HALL, M. and MARJORAM, D. T. E., 1973, Mathematics in a Changing World (Heinemann).
15. HULL, J., MAPES, J. and WHEELER, B., 1976, Model Building Techniques for Management (Saxon House).
16. LANCASTER, P., 1976, Mathematical Models of the Real World (Prentice Hall).
17. LIGHTHILL, J., 1978, Newer Uses of Mathematics (Penguin).
18. McDONALD, T., 1974, Mathematical Methods for Social and Management Scientists (Houghton Mifflin).
19. NOTARI, R. E., 1975, Biopharmaceutics and Pharmacokinetics (Dekker).
20. MAKI, D. P. and THOMPSON, M., 1973, Mathematical Models and Applications (Prentice Hall).
21. MURDICK, R., 1971, Mathematical Models in Marketing (Intext)
22. OPEN UNIVERSITY 1977, Course TM281 "Modelling by Mathematics" (O.U. Press).
23. TEICHROEW, D., 1964, Introduction to Management Science Deterministic Models (Wiley).

24. WALFORD, R., 1959, Games in Geography (Longman).
25. WILSON, A. G. and KIRKBY, M. J., 1975, Mathematics for Geographers and Planners (Clarendon).

Below are the appropriate references for the examples in this booklet.

2 – King Arthur's Round Table
References 3, 7 and 20.

3 – Traffic Flow
References 13, 22 and The Highway Code.

4 – Clunk Click
References 5 and 22.

5 – Optimum Order Quantity
References 3, 15, 21 and 23.

6 – Single Species Population Models
References 1, 4, 5, 7, 12, 13 and 14.

7 – Launching a Satellite
References 1 and 8.

8 – "Save It" – Water Heating and Cooling
References 5 and 22.